Transition metal catalysts within protein scaffolds– three case studies on the development and engineering of artificial metalloenzymes

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Jingming Zhao

Aus Anyang, China

Basel, 2018

Genehmigt von der Philophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Thomas R. Ward

Prof. Dr. Dennis G. Gillingham

Basel, den 27. 03. 2018

Prof. Dr. Martin Spiess

Dekan

Acknowledgment

This thesis is a summary of several years' brain storming and hard lab work. It would be not possible without the support and guidance of so many kind and great masters, to whom I would like to express here my sincere appreciation.

First of all, I would like to thank Prof. Thomas R. Ward for his full support, guidance, enormous patience and stimulation to me during my whole PhD study. He taught me not only science but also showed me the great courage to face difficult and challenging issues and strive to find out solutions, which brings the research to a very high level and quality.

I am grateful to Prof. Dennis G. Gillingham for being a co-referee to evaluate this thesis and his great help on the dirhodium project to me. I thank Prof. Christof Sparr for being the chairman of my PhD defense.

I greatly thank Prof. Wolf-Dietrich Woggon for his inspiring discussion and assistance to me in the synthesis of different types of dirhodium complexes.

My great appreciation also to Dr. Daniel Häussinger for endless 600 MHz 2D NMR measurement of all my final biotinylated dirhodium complexes and special water suppression NMR measurement to track reaction intermediates.

I specially thank Dr. Valentin Köhler for all scientific help and analytical instrument assistance. It is not possible for me to start enjoying Basel life at the beginning without all his private help.

Furthermore I should thank Dr. Raphael Reuter, Dr. Anna Kajetanowicz Dr. Anamitra Chatterjee, Dr. Yoann Cotelle, Dr. Isabel Teresa Alt, Dr. Xingwei Guo, Dr. Qi Zhang, Dr. Lei Yang and Ke-Feng Zhang for organic chemistry discussion and assistance.

I thank a large amount of great biologists in the lab who save my life a lot: Dr. Hendrik Mallin for primer design and mutant library construction for the directed evolution; Dr. Michela Maria Pellizzoni for providing the chimeric loop protein construct; Dr. Johannes Rebelein for the restriction digest, PCR and ligation assistance, and Rosetta design for the streptavidin-FPD structure; Juliane Klehr, Dr. Shuke Wu, Dr. Yi Zhou, Dr. Emeline Sautron, Dr. Ryan Peterson, Dr. Cangsong Liao and Dr. Aping Niu for general protein expression and purification help and teaching; A big thank you to Dr. Christian Trindler for his patient supervision on periplasm screening assays.

My appreciation also to Dr. Tillmann Heinisch for solving the X-ray crystal structures and directed evolution discussions; Maxime Barnet for his professional docking simulation performance; I thank all the catalysis experts for all their wonderful ideas: Dr. Yasunori Okamoto, Dr. Fabian Schwizer, Dr. Vincent Lebrun, Dr. Joan Serrano Plana, Dr. Marc Dürrenberger, Dr. Sascha Keller and Dr. Zhe Liu.

I thank all the current group young students who help to generate very friendly and social lab environment: Martina Riba Hestericova, Valerio Sabatino, Miriam Kuhn, Jaicy Vallapurackal, Fadri Christoffel, Jonas Schätti and Boris Lozhkin.

I also thank Mrs. Isa Worni, Mrs. Esther Stalder and Mrs. Beatrice Erismann for all their effective administrative help, the Werkstatt team's great contribution to keep the whole department running efficiently.

Finally I am grateful to my mom, dad and other family members in China. They offer me their deep love and long term support and encouragement during my PhD study.

And a big thank you, my dear wife Dr. Zi Liu, for all your support, understanding and deep love all the time. Your excellent cooking survived me during the hardest time in these years. I enjoyed all our wonderful travelling around Europe, the skiing time, hiking and creative scientific discussions…

Summary

Transition metal catalysis provides versatile methods for the preparation of synthetically useful compounds. Some of these transformations can also be realized by enzymatic catalysis at much higher activity and selectivity. The research field of artificial metalloenzymes (ArMs) lies at the interface of traditional transition metal catalysis, organocatalysis and enzymatic catalysis. In this line of research, new-to-nature transition metal complexes are incorporated into suitable protein scaffolds to explore the reactivity and selectivity of the resulting constructs, thereby greatly expanding the repertoire of enzymatic catalysts. To match the high efficiency and selectivity displayed by enzymatic catalysts, ArMs need to be optimized via effective engineering methods such as directed evolution.

Within the scope of this thesis, complexes of the transition metal elements Ru, Rh, and Ir were combined with mutants of human carbonic anhydrase II (hCA II), and streptavidin from *Steptomyces avidinii*, respectively. Specifically, three types of ArMs were developed : i) an artificial metathesase based on human carbonic anhydrase II mutants, ii) a dirhodium carbenoid transferase and iii) artificial transfer hydrogenase, the latter two based on streptavidin (Sav) mutants as the protein scaffold.

The first subsection of chapter 1 provides a general introduction on ArMs, including anchoring and engineering strategies and presents their synthetic applications. A particular focus is set on the achievements of ArMs in the past two years. The further subsections of chapter 1 detail important background information for the constructs under investigation.

Chapter 2 describes the incorporation of arylsulfonamide anchored Hoveyda-Grubbs Ru-catalysts into human carbonic anhydrase II to create an ArM for ring-closing metathesis. The binding affinity of a novel Ru-cofactor for hCA II scaffold was determined. The catalytic activity of the ArMs was evaluated with different olefin substrates under optimized conditions. The best ArM/substrate pair was employed in test reactions under pH-neutral conditions at a low catalyst concentration (10 μM). Such studies are an important prerequisite for the application of ArMs *in vivo*.

In chapter 3, the development of a dirhodium based artificial carbenoid transferase for intermolecular cyclopropanation and C-H insertion reactions is outlined. ArMs which combine either bidentate or monodentate dirhodium carboxylate complexes with streptavidin were compared in terms of their carbene transfer activity. Docking simulations of the biotinylated dirhodium catalysts within streptavidin predicted the position of the rhodium center in respect to the biotin-binding vestibule. Biocompatibility of the dirhodium complex was tested by performing the dirhodium ArM catalyzed cyclopropanation reaction in the presence of whole *E. coli* cells.

Chapter 4 summarizes a study on the directed evolution of an iridium based artificial transfer hydrogenase. The starting point was a computed structure of a chimeric streptavidin scaffold containing an inserted loop partially shielding the active site, which was predicted by Rosetta design. Site-directed mutagenesis at four residues was performed to build up a mutant library for screening. A self-immolative substrate which releases a fluorescent product upon reduction was developed and used in the evolution process to increase the screening throughput. Around 1000 clones were screened with

the hits sequenced. The promising mutants were purified and their performance was investigated for imine reduction *in vitro*.

Table of contents

Preface

All schemes, figures, tables and compounds of each chapter in this thesis are numbered independently. All experimental details, references for individual chapters are collected at the end of each chapter.

Chapter 1 | Introduction

1.1 Artificial metalloenzymes

Artificial metalloenzymes (ArMs hereafter) incorporate a catalytic metallocofactor into a protein scaffold. Such systems have attracted both chemists and biologists for many years.[1] By identifying various metal complexes stable and active in an aqueous environment, chemists are able to introduce new-to-nature reactions in a biological setting.[2] With the help of protein engineering and mutagenesis techniques, the homogeneous catalysts' activity can be efficiently improved using high throughput techniques. Thanks to protein design strategies, ArMs' structures and functions can be effectively engineered and readily optimized.[3]

There are four well established approaches to anchor synthetic catalysts to protein scaffolds: covalent linking (using a covalent bond), supramolecular anchoring (main contributions from the Ward group, biotin-streptavidin technology as the most commonly employed tool), dative anchoring (direct coordination to the metal center) and metal substitution (native metal ion is replaced by a new metal ion). (Figure 1)

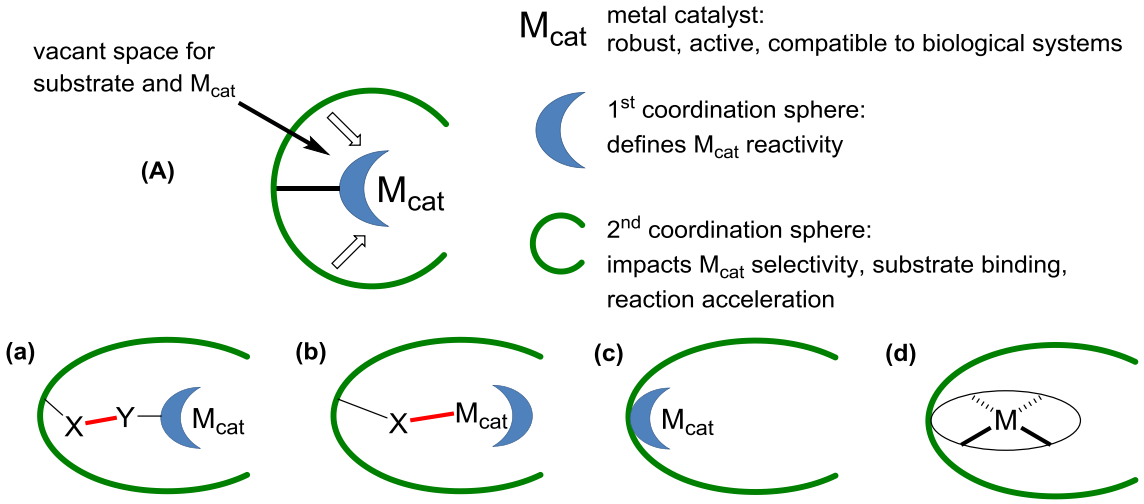


Figure 1. (A) General structure of an artificial metalloenzyme. An abiotic cofactor is localized within the protein scaffold via (a) covalent linking, (b) dative anchoring, (c) supramolecular interaction and (d) metal substitution.

In recent years, outstanding research papers[4–8] and reviews[9–15] about ArMs with novel cofactors and protein scaffolds have been reported, mainly focused on catalytic activity, mechanistic studies, structure design and engineering. I summarize below the most recent reports on ArMs after the publication of the comprehensive review by the Ward group and Lewis group.[1]

Myoglobin is currently a very popular scaffold to assemble an ArM to catalyze many transformations. As it is naturally evolved for binding dioxygen, researchers have explored its catalytic uses by either substituting the metal center in the heme or modifying the porphyrin structure.

The Hayashi group provided detailed mechanistic study of a manganese porphycene reconstituted myoglobin (rMb) for C-H hydroxylation.[16] At pH 8.5, a high-resolution crystal structure of rMb highlighted the stable ligation of His93 to Mn center and no alternative conformation of His64, which are critical for the promotion of C-H hydroxylation. A single turnover reaction and EPR spectroscopy allowed the characterization of the Mn^{V}-oxo species, the first characterized high-valent species in a chemically engineered hemoprotein. (Figure 2)

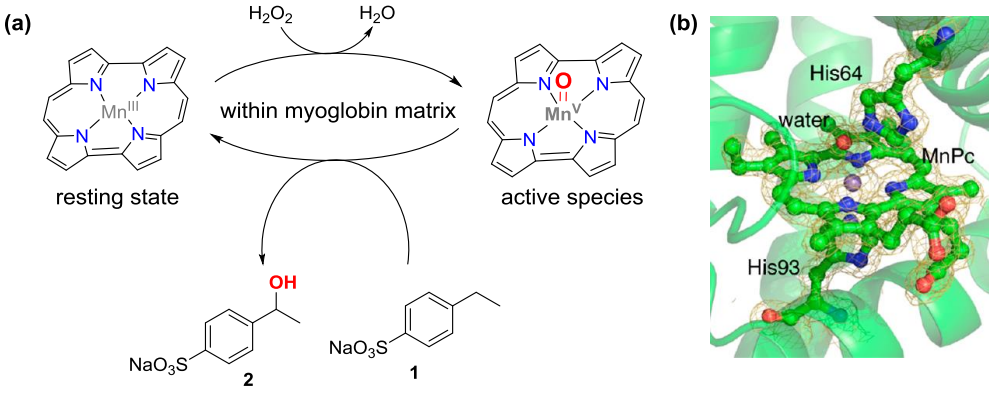
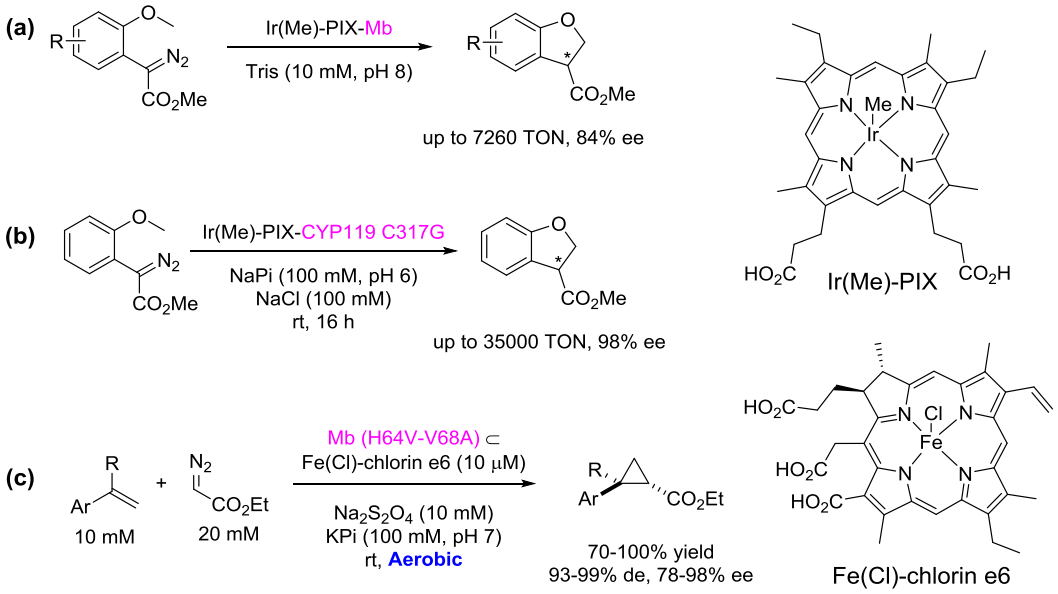


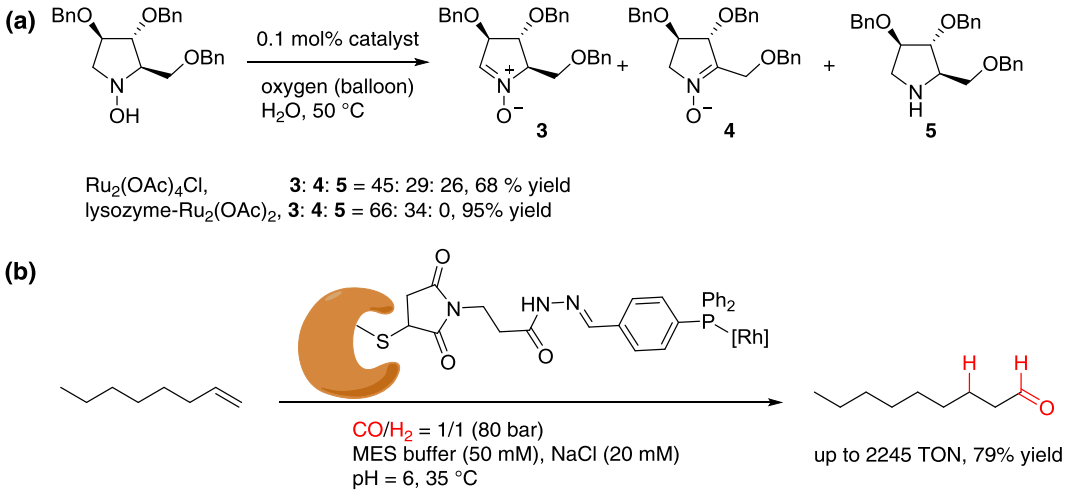
Figure 2. (a) Proposed reaction mechanism of C-H hydroxylation of water soluble substrate sodium 4-ethylbenzenesulfonate within rMb. (b) Crystal structure of rMb at pH 8.5 with His64, His93 and water molecule around MnPc center.

More recently, the Hartwig group introduced an abiological metal substitution concept for a non-natural C-H insertion reaction, using Ir(Me)-PIX substituted myoglobin.[17] (Scheme 1a) Furthermore, they also incorporated the iridium porphyrin into P450 enzyme CYP119. This repurposed enzyme displayed kinetic parameters comparated to natural enzymes for C-H insertion.[18] (Scheme 1b) The Fasan group reported myoglobin catalyzed carbene-mediated cyclopropanation under aerobic conditions with high catalytic efficiency (up to 6970 TON) and stereoselectivity (up to 99% de and ee).[19] They achieved this by substituting the natural heme cofactor with iron-chlorin e6 complex. (Scheme 1c)



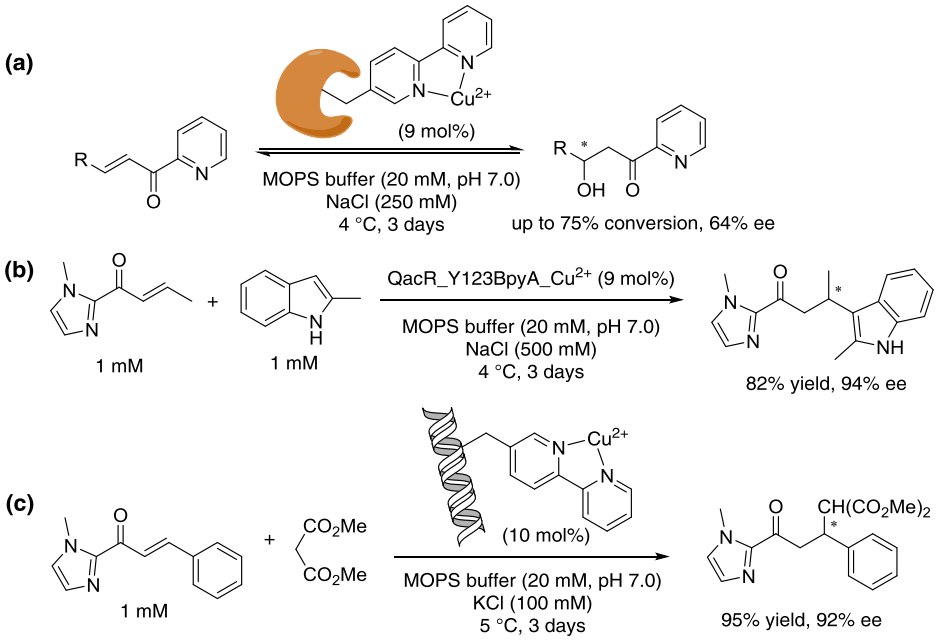
Scheme 1. A metal substitution strategy was used to assemble new efficient and selective ArMs. Ir-reconstituted heme within myoglobin (a) and P450 scaffolds (b) catalyze intramolecular C-H insertion. (c) Fe(Cl)-chlorin e6 incorporated myoglobin catalyzes stereoselective olefin cyclopropanation.

There are also examples of ArMs as catalysts in non-natural compounds synthesis. Goti and coworkers reported a lysozyme-$Ru_2(OAc)_2$ complex that catalyzes the aerobic oxidation of hydroxylamines to nitrones. The $[Ru_2(\mu\text{-}OAc)_2\text{-}(H_2O)_2]^{3+}$ moiety is hypothesized to bind to two exposed aspartate groups, affording different chemoselectivity for the oxidation, compared to $Ru_2(OAc)_4Cl$.[20] (Scheme 2a) The Kamer group reported an artifical hydroformylase by combining rhodium and a lipid-binding protein scaffold through robust site-specific phosphine bioconjugation. The newly formed ArM displayed remarkable activity and selectivity for the hydroformylation of linear aldehydes.[21] (Scheme 2b)



Scheme 2. (a) The selective diruthenium incorporation in lysozyme affords an ArM that catalyzes the oxidation of hydroxylamines. (b) Highly efficient and selective rhodium hydroformylase for the synthesis of linear long-chain aldehydes.

The Roelfes group reported their effort on designing an enantioselective artificial metallo-hydratase based on the transcriptional repressor lactococcal multidrug resistance regulator (LmrR).[22] The non-canonical amino acid (2, 2'-bipyridin-5yl)alanine (BpyA) residue introduced in the LmrR scaffold binds the catalytic Cu^{2+} ion for enantioselective hydration of the C=C bond of enones. (Scheme 3a) The agreement between computational- and experimental results highlighted the power of computational design for the creation of novel artificial metallohydratases. The same group also incorporated a Cu^{2+} ion into other types of Multidrug resistance regulators (MDRs) as protein scaffolds. The resulting ArMs catalyzed the enantioselective Friedel-Crafts alkylation with up to 94 % ee. (Scheme 3b) The large hydrophobic and promiscuous binding pocket facilitates the ArMs' design.[23] The Jäschke group reported Cu^{2+} ion bound G-quadruplex DNA-based ArMs for enantioselective Michael addition in water.[24] (Scheme 3c) Cu^{2+} was covalently anchored to the bipyridyl liangds in the G-quadruplex DNA. CD spectra indicated the limited influence of the linker length on the quadruplex structure.



Scheme 3. (a) Enantioselective hydration of enones catalyzed by Cu^{2+} incorporated LmrR mutants. (b) Cu^{2+} catalyzed vinylogous Friedel-Crafts alkylation using MDR family as the protein scaffolds. (c) G-quadruplex DNA-Cu^{2+} complex catalyzed asymmetric Michael addition.

ArMs also draw inorganic chemists' interest for structural, spectroscopic and mechanism studies. Fujieta and coauthors reported an artificial osmium peroxygenase relying on a metal-substitution strategy of the TM1459 cupin superfamily protein as scaffold. The Os ion binds four histidines in the active site, resulting in an octahedral coordination geometry and contributing to stabilize the protein's quaternary structure. The reconstituted ArM catalysed the dihydroxylation of alkenes, affording diols up to 9100 TON. Site-directed mutagenesis of this osmium peroxygenase resulted in a threefold catalytic improvement.[25] Wang and coworkers reviewed recent progress on improving ArMs' activity

by optimizing electron-transfer process. Strategies to achieve this include: tuning cofactor's redox potential, tailoring the cofactor's spatial arrangement, introducing a protein redox partner and etc.[12]

In summary, ArMs are becoming a hot research topic. ArMs attract organic chemists to develop new catalytic reactions, inorganic chemists to resolve new metal-protein structures. ArMs also appeal to enzymologists to explore new protein functions, biologists to discovery cascade reactions with natural enzymes and computational scientists to model substrate-protein interactions, predict transition state and design new types of ArMs.

1.2 Human carbonic anhydrase II

1.2.1 Basic knowledge of human carbonic anhydrase II

Human carbonic anhydrase II (hCA II) is a monomeric, single-chain metalloenzyme (Figure 3a) that catalyzes the reversible hydration of CO_2 into bicarbonate (HCO_3^-) and a proton (H^+) at an extremely high rate, with a k_{cat}/K_m of 1.5 X 10^8 $M^{-1}s^{-1}$ and a k_{cat} of 1.4 X 10^6 s^{-1} (eq. 1, Figure 4).[26] The active site of hCA II comprises a catalytic Zn^{2+} ion coordinated to three imidazole groups of histidines (His94, His96 and His119) and one OH^- ion/water molecule. Hydrogen bonding interactions between the zinc-bound water and the hydroxyl moiety of Thr199, which is in turn bridged to the carboxylate moiety of Glu106, enhance the nucleophilicity of the zinc-bound water molecule, and further orient CO_2 substrate in a favorable location for nucleophilic attack. pKa of the coordinated H_2O to Zn^{2+} is lowered to ~6.8 (compared with the pKa of bulk H_2O ~14), which sufficiently facilitates the Zn^{2+}-OH^- formation at physiological pH.[27] (Scheme 4)

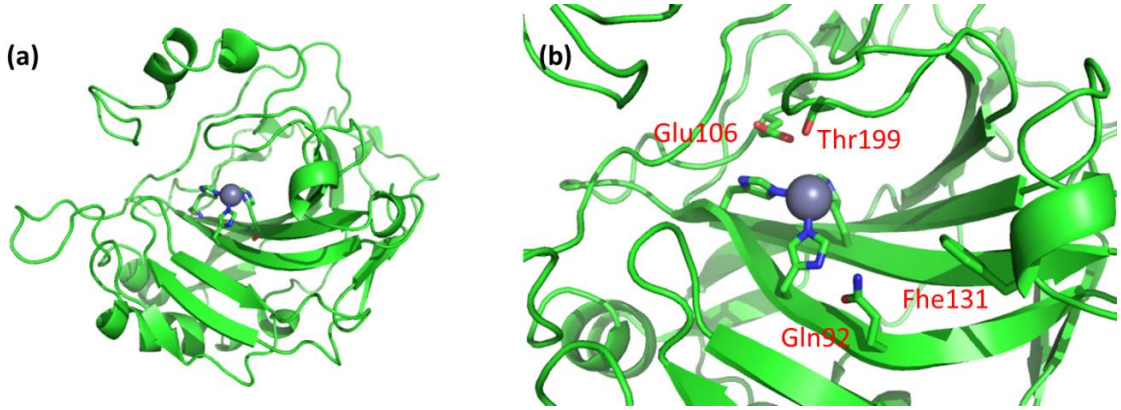
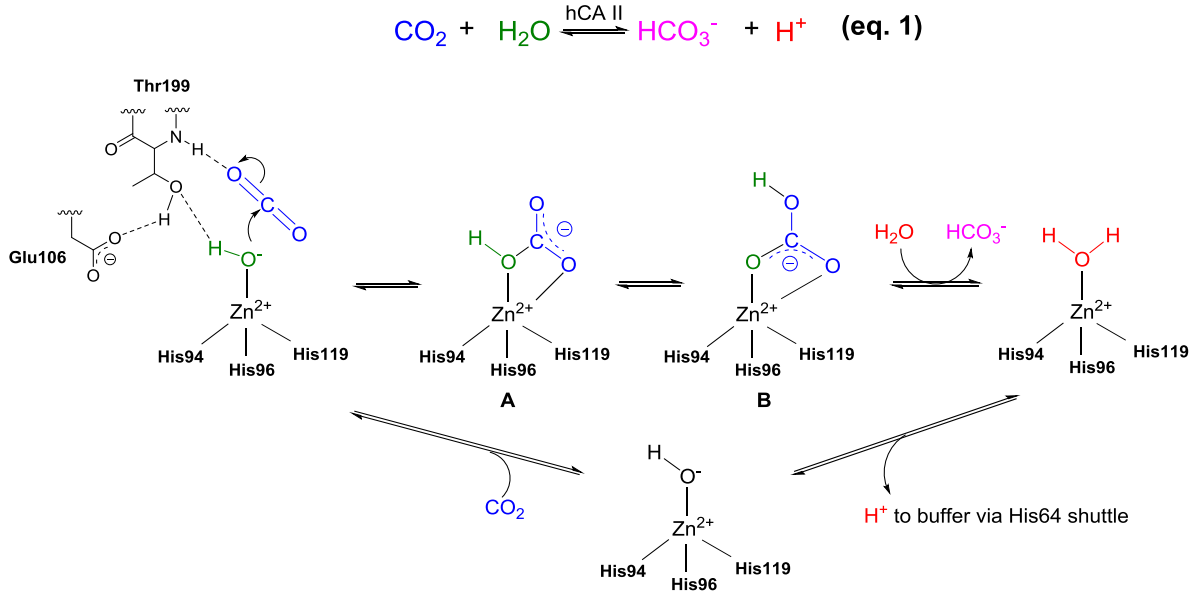


Figure 3. (a) X-ray structure of native hCA II (PDB: 2CBA) with a Zn^{2+} ion (gray) coordinated to the three histidine residues. (b) Close-up-view of the critical residues involved in the substrate or inhibitor binding in the active site.



Scheme 4. Mechanism of the hCA II-catalyzed CO_2 hydration. Nucleophilic attack of the zinc-bound OH^- to CO_2 forms the zinc-bound bicarbonate complex **A** which isomerizes to **B**. After the exchange with H_2O in intermediate **B**, the bicarbonate product is released and H^+ is transferred from zinc-bound water to regenerate the zinc-hydroxide species via His64.

Aryl sulfonamide-containing compounds are common inhibitors (K_d = 200-1500 nM) for hCA II, in which the residues Gln92, Glu106, Phe131 and Thr199 are involved in their tight binding (Figure 3b). The deprotonated sulfonamide is coordinated to Zn^{2+} in the active site. The NH group also interacts with Thr199 through a hydrogen bonding, which in turn interacts with the carboxylate group of Glu106 by another hydrogen bond. One oxygen atom of the SO_2NH moiety binds with the backbone NH group of Thr199 through H-bond, with another oxygen atom weakly coordinated to Zn^{2+} (Figure 4a). Variation of the R group leads to different interactions with the hydrophobic surface, significantly affecting the inhibitor's binding affinity: for benzene sulfonamide K_d vary between 200-1500 nM and for acetazolamide K_d = 12 nM (Figure 4b and 4c).[28,29]

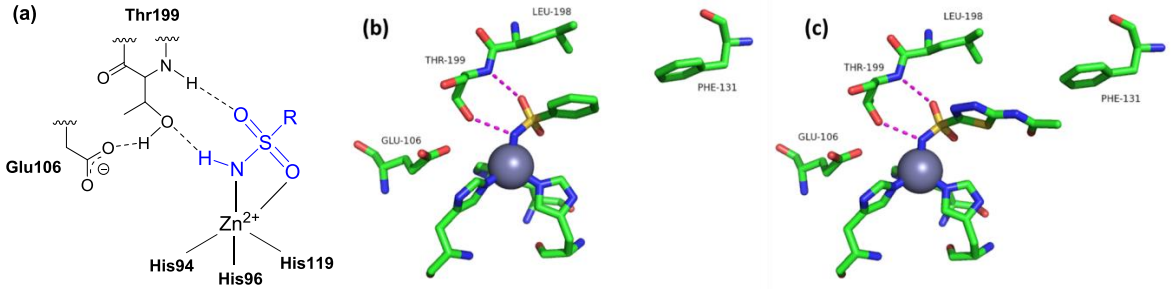


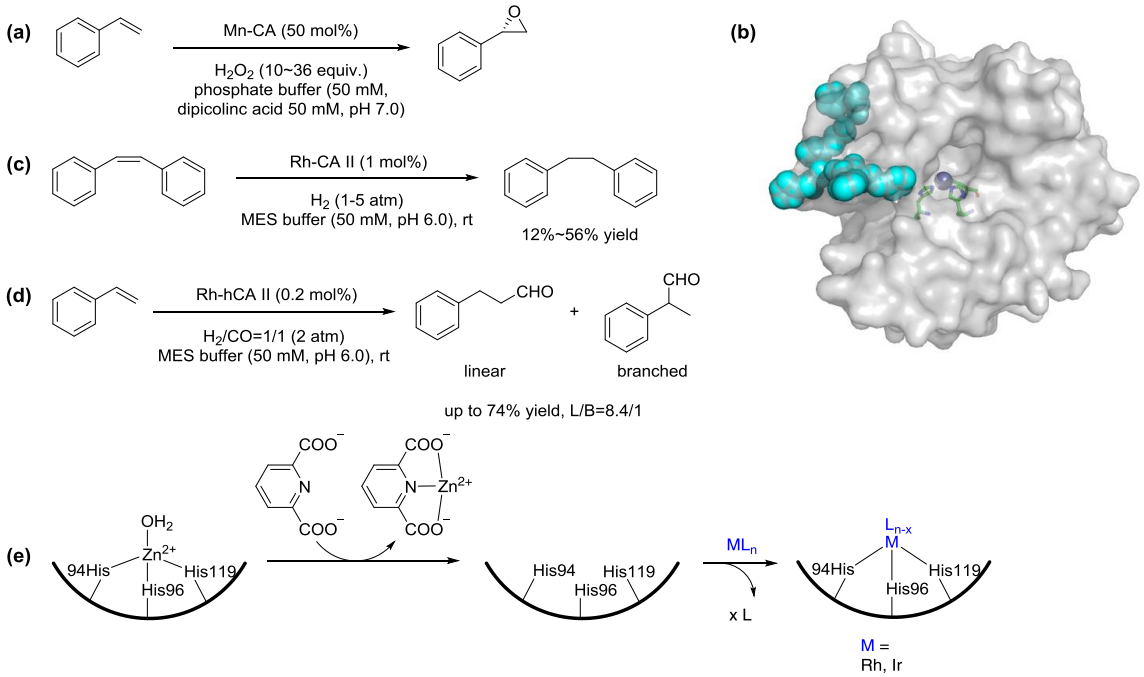
Figure 4. (a) hCA II inhibition mechanism by sulfonamide inhibitors. (b) Structure of benzenesulfonamide bound to hCA II (PDB: 2WEJ), magenta dashed lines represent hydrogen bond. (c) Structure of acetazolamide (AZM) bound to hCA II (PDB: 3HS4, the other two surface bound AZM molecules are not shown)[30].

1.2.2 Human carbonic anhydrase II as a scaffold for the creation of ArMs

Human carbonic anhydrase II has received significant attention to construct ArMs in recent years. Two strategies have been pursued: (a) substitution of the native Zn^{2+} with other transition metals to reconstitute the whole scaffold; (b) incorporation of transition metal catalysts bearing arylsulfonamide achors to install new catalytic activities.

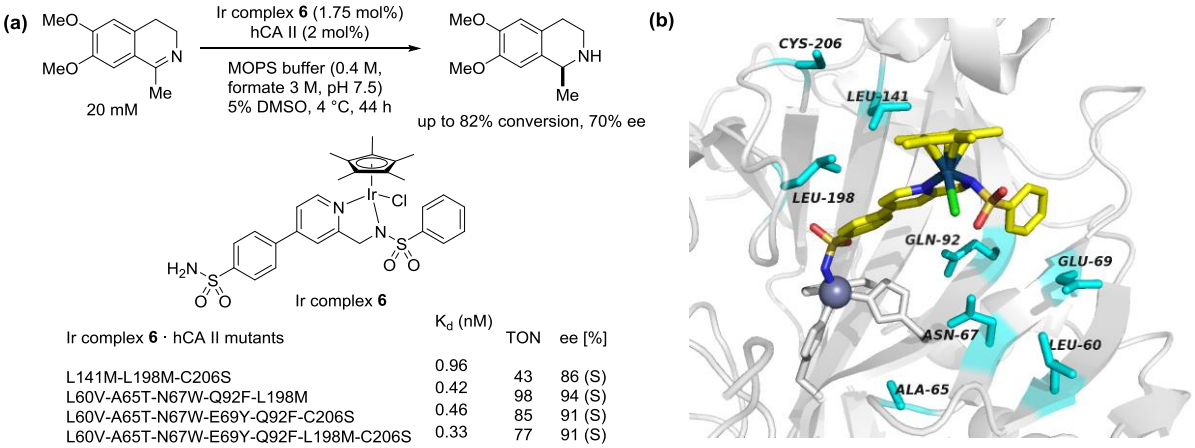
Soumillion and coworkers reported the first Mn-substituted carbonic anhydrase for enantioselective epoxidation of olefins.[31] Although the reconstituted Mn-CA displayed lower affinity (micromolar range) for the manganese ion, it achieved up to 57% yield and 52% ee for asymmetric epoxidation of styrene, higher than the free manganese (Scheme 5a, 5b). The Kazlauskas group succeeded in replacing the active-site Zn^{2+} cofactor in carbonic anhydrase with $[Rh(cod)_2]BF_4$ and $[Rh(acac)(CO)_2]$ to achieve the stereoselective hydrogenation of stilbene[32] and the regioselective hydroformylation of styrene[33], respectively (Scheme 6c, 6d). To minimize the non-specific binding of Rh(I) to the CA

surface, the authors removed the surface histidine residues by chemical modification and site-directed mutagenesis. Side products (isomerized *trans*-stilbene and branched aldehyde) were produced by the rhodium located outside of the active site. Lee and coworkers developed the first iridium substituted carbonic anhydrase for water oxidation under mild and neutral conditions.[34] The newly formed Ir-carbonic anhydrase complex displayed comparable oxygen-evolving activity to traditional Ir-based small molecule catalysts. The Hartwig group reported the preparation and characterization of rhodium and iridium-substituted carbonic anhydrase.[35] (Scheme 5e) Advanced characterization methods were used in the binding studies, such as colorimetric assay (to quantify the metal occupancy at the native metal-binding site), ^{15}N-^{1}H NMR spectroscopy (to establish the amino acids to which the metal is bound), UV-vis spectroscopy (to detect the species during the substitution reaction) and etc.



Scheme 5. (a) Mn-CA catalyzes the enantioselective styrene epoxidation. (b) Surface-exposed histidines (His3, His4, His10, His15, His17 are highlighted in cyan, PDB: 1G1D) were pretreated with diethyl pyrocarbonate to minimize metal-binding to surface histidines. (c) Rh(I)-CA II catalyzed hydrogenation. (d) Rh(I)-hCA II catalyzed hydroformylation. (e) Standard method used to remove the native Zn^{2+} ion from carbonic anhydrase by novel metals.

The Ward group first developed an artificial transfer hydrogenase by incorporating arylsulfonamide-bearing iridium complexes [η^5-Cp*Ir(pico)Cl] with hCA II scaffold (Scheme 6a). After both chemical and genetic optimization, complex **6** proved to be the best and achieved up to 82% conversion and 70% ee for the transfer hydrogenation of salsolidine precursor (Scheme 6a).[36] Based on the crystal structure of complex **6** · hCA II WT and Rosetta design protein engineering, eight positions in the active site were selected for further genetic optimization (Scheme 6b). The newly designed mutants displayed significant improvement for binding affinity (46-64 fold), activity (TON from 9 to 98) and enantioselectivity (ee from 70% to 96%).[37]



Ir complex **6** · hCA II mutants	K_d (nM)	TON	ee [%]
L141M-L198M-C206S	0.96	43	86 (S)
L60V-A65T-N67W-Q92F-L198M	0.42	98	94 (S)
L60V-A65T-N67W-E69Y-Q92F-C206S	0.46	85	91 (S)
L60V-A65T-N67W-E69Y-Q92F-L198M-C206S	0.33	77	91 (S)

Scheme 6. (a) Ir complex **6** · hCA II catalyzes transfer hydrogenation of the salsolidine precursor and K_d values for the new designed mutants. (b) X-ray crystal structure of Ir complex **6** · hCA II WT (PDB: 3ZP9) resulting from the Rosetta Design; the mutations suggested by the computational design are highlighted in cyan.[37]

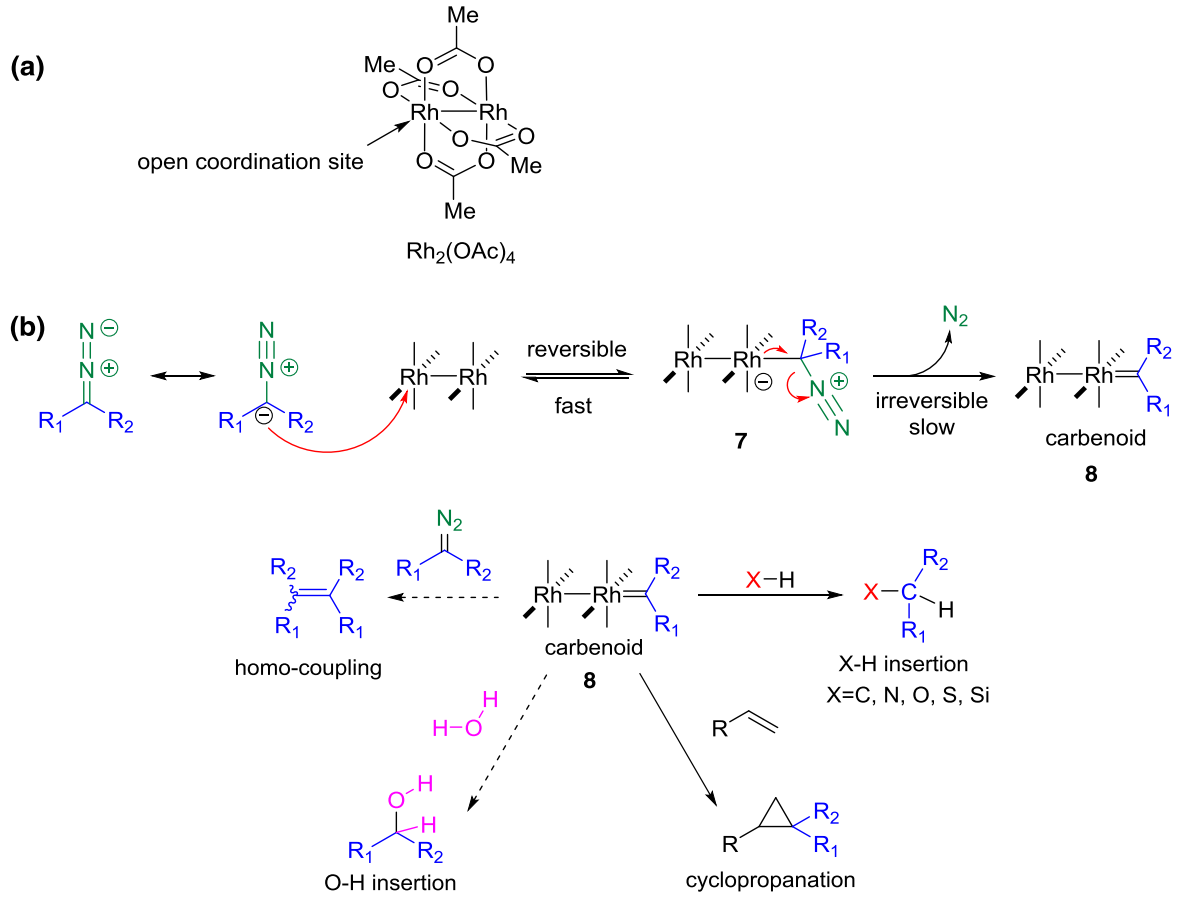
1.2.3 Summary

Human carbonic anhydrase has been thoroughly investigated concerning its thermostability[38], protein-ligand binding, inhibitor development[39], therapeutic applications[40,41] and etc. However, it has not been widely studied as a protein scaffold to construct ArMs for new-to-nature reactions. In the course of this thesis, I have contributed to develop a biocompatible artificial metathesase using hCA II as host protein (see chapter 2).

1.3 Bioconjugation of dirhodium complex for the creation of ArMs

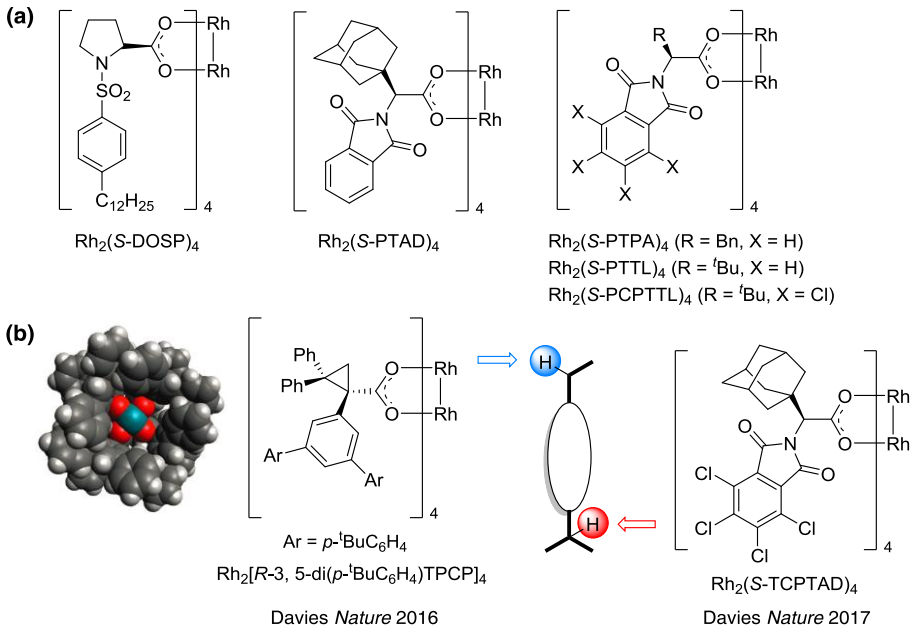
1.3.1 Dirhodium tetracarboxylate complexes

Dirhodium (II) tetracarboxylate complexes have proven to be stable, robust and highly active catalysts for carbenoid transfer reactions, such as cyclopropanation, X-H insertion (X=C, N, O, S, Si)[42], arylation of boronic acids[43], dearomatization of isoquinolinium[44] and so on. They are also used as building blocks for self-assembly[45] and immobilization on solid support[46] for heterogeneous catalysis. Dirhodium (II) acetate is the simplest dirhodium complex (Scheme 7a). The rhodium atom has 16 electrons with an open coordination site to accept two electrons from a carbene source (such as diazo compounds) via axial coordination, resulting in an eighteen electron complex **7**. After the irreversible extrusion of N_2 from intermediate **7**, Rh(II) carbenoid **8** is formed (Scheme 7b). This carbenoid forming step is considered to be the rate-limiting step in the catalytic cycle.[47] The Fürstner group has reported the X-ray crystal structure of donor/acceptor donor/donor dirhodium carbene in the solid state.[48] Starting from **8**, various synthetically useful transformations have been achieved. In aqueous media, the homocoupling of the diazo compound and the O-H insertion environment need to be minimized (Scheme 7b).

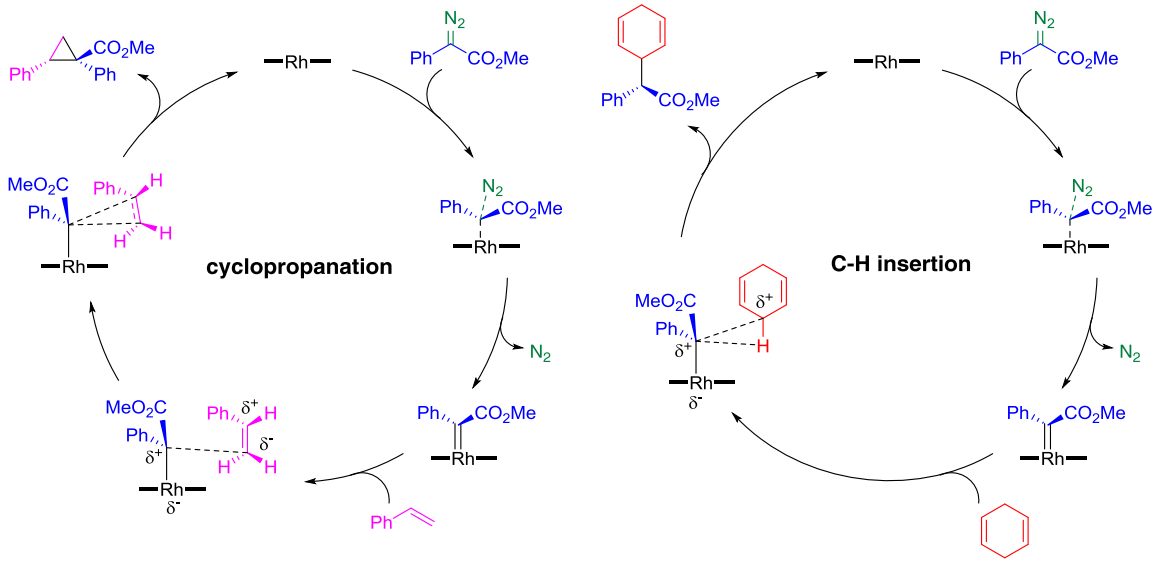


Scheme 7. (a) Structure of dirhodium acetate. (b) Proposed mechanism of Rh(II) carbenoid formation and various carbene transfer reactions.

In recent years, highly efficient dirhodium catalysts have been developed by replacing the acetate ligands with other bulky carboxylate ligands for asymmetric synthesis, generating excellent diastereo and enantioselectivity. The Davies group has contributed most on developing novel chiral dirhodium catalysts for cyclopropanation and C-H functionalization. These include Rh_2(*S*-DOSP)$_4$, Rh_2(*S*-PTAD)$_4$ and so on (Scheme 8a).[49] Very recently, they reported on even more bulky dirhodium catalysts for site and stereo selective insertion of unactivated C-H bonds. Secondary and tertiary C-H bonds were selectively functionalized by catalysts control.[50,51] (Scheme 8b)



Scheme 8. (a) Examples of widely used chiral dirhodium tetracarboxylate catalysts. (b) Site-selective and catalyst-controlled C-H functionalization.



Scheme 9. Proposed mechanisms of the dirhodium catalyzed cyclopropanation and C-H insertion.

The mechanism of dirhodium catalyzed cyclopropanation and C-H insertion are presented in scheme 9. The rhodium carbenoid formation is always rate determining. In the cyclopropanation cycle, forming a

cyclopropane ring from an electrophilic carbene and a nucleophilic olefin is generally thought to occur in a nonsynchronous manner. The carbene initially reacts with more electron-rich olefinic carbon.[52] Also in the C-H insertion cycle, sites that stabilize the positive charge build-up are favored for C-H functionalization, forming a three-center, two-electron transition state.[53] To obtain good diastereoselectivity, the highly reactive and electrophilic rhodium carbenoid intermediate needs to be stabilized by diazoacetate substrates. Donor/acceptor rhodium cabenoids display enhanced sensitivity towards both steric and electronic factors. In general, 2° C-H bonds are most favored for carbene insertion because 2° site can best stabilize a positive charge without being too sterically crowded. Although 3° C-H bonds are more activated than 2° C-H bonds on an electronic basis, they can only be functionalized if they are not too crowded.[54]

With chiral and bulky tetracarboxylate ligands, the corresponding dirhodium catalysts display exceptional enantioselectivity by blocking selected orientations of the incoming substrates. To illustrate the versatility of such systems, let's focus on Rh_2(*S*-DOSP)$_4$, a highly stereoselective and versatile catalyst for both cyclopropanation and C-H insertion. (Figure 5a) There are four distinct vectors along which a substrate may approach the rhodium carbene moiety. The D_2-symmetric tetraprolinate DOSP ligands can block two pathways efficiently. The third pathway is blocked by the steric bulk of the ester group from the carbene on the back face. This leaves only one vector for the alkene substrate to approach the catalyst. (Figure 5b) The bulkier the carboxylate ligands, the higher enantioselectivity for the reaction.

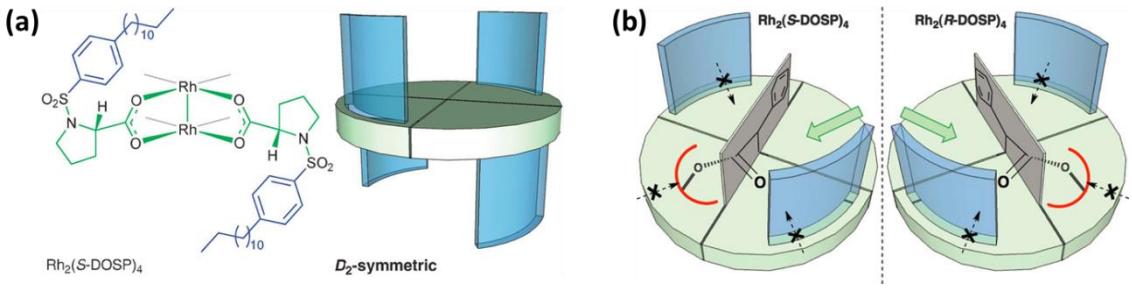
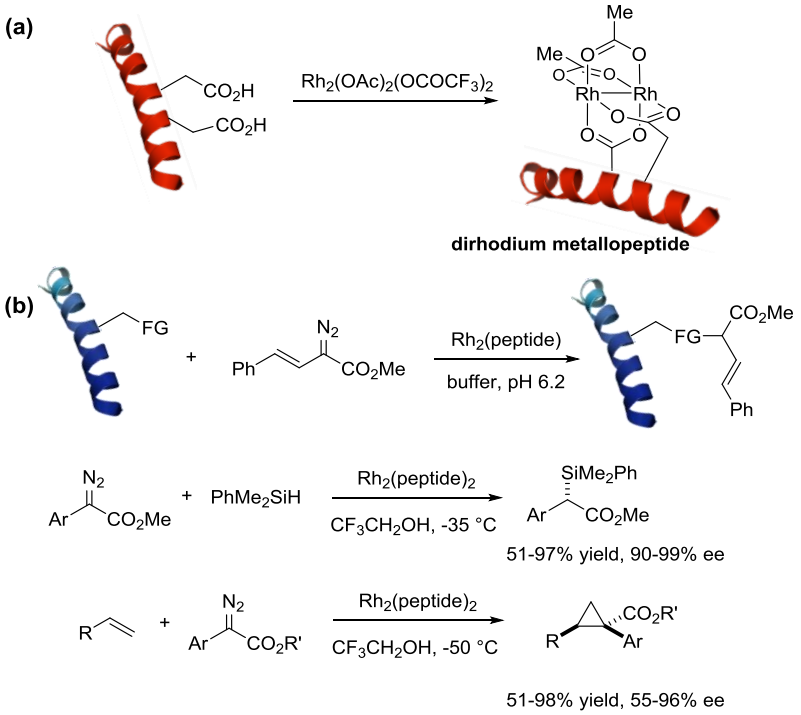


Figure 5. (a) Structure of D_2-symmetric Rh_2(*S*-DOSP)$_4$ catalyst, the blue disc-like shapes represent tetraprolinate ligands in quadrants-separated configurations. (b) Modelled top view of steric factors influencing the facial selectivity of the Rh_2(*S*-DOSP)$_4$ stabilized carbene.

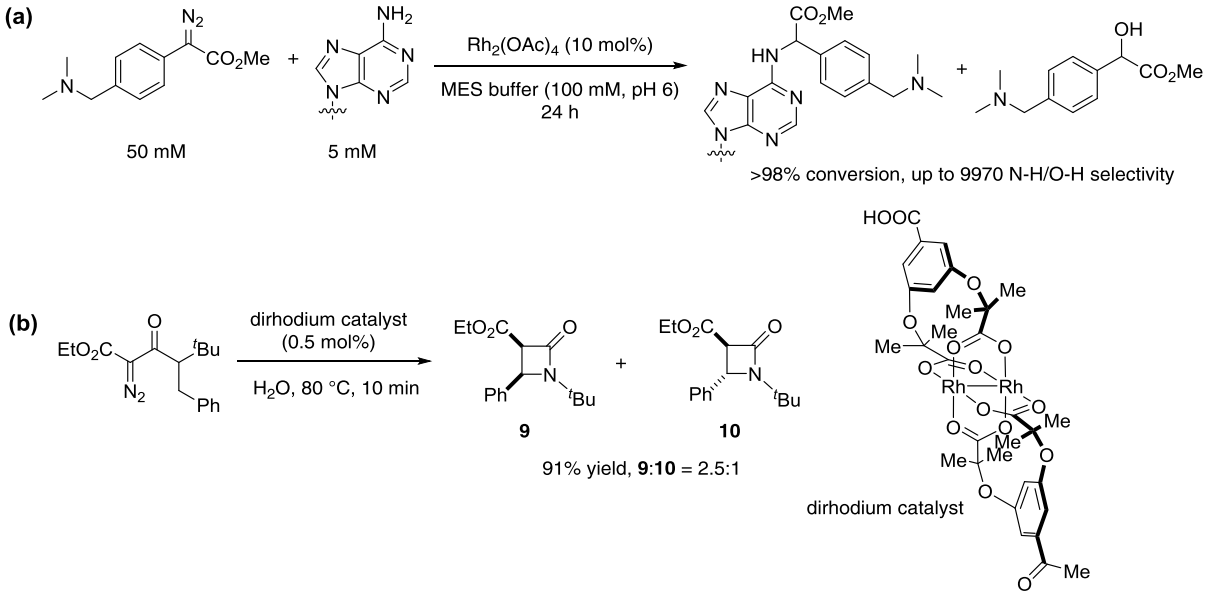
1.3.2 Introducing dirhodium moieties as biocatalysts for carbene transfer reactions

Thanks to their robustness and great activity, dirhodium complexes also draw attention from the chemical biology community. Many useful transformations catalyzed by dirhodium catalysts have been reported under biological conditions. Ball and coauthors first developed dirhodium metallopeptides by anchoring the dirhodium mioety to two glutamate side chains from a peptide.[55] These dirhodium metallopeptides catalyze: i) the modification of tryptophan, ii) enantioselective carbenoid inertion into Si-H bond[56] and iii) asymmetric cyclopropanation[57]. (Scheme 10) More

importantly, these catalysts proved to be active in *E. coli* lysates[58], paving the way for dirhodium catalyzed transformations in *E. coli* cells. Instead of using traditional dirhodium catalysts with chiral carboxylate ligands, the Ball group first introduced biomolecule peptides as the chirality source. The Gillingham group reported dirhodium catalyzed alkylation of DNAs and RNAs with good conversions and excellent selectivity.[59] (Scheme 11a) Later, the same group developed various mono- and bis-substituted dirhodium complexes with different modular carboxylate ligands.[60] (Scheme 11b) These dirhodium complexes catalyzed intramolecular C-H insertion with high efficiency even in aqueous environment. Certain dirhodium complexes showed great potential biological applications, such as specific localization to DNA in living cells.



Scheme 10. (a) Formation of dirhodium metallopeptide. (b) Dirhodium metallopeptides catalyze carbene insertion transformations.



Scheme 11. (a) Dirhodium acetate catalyzed alkylation of DNAs and RNAs. (b) Aqueous intramolecular C-H insertion catalyzed by newly developed dirhodium complex.

The Lewis group first created a dirhodium-based ArM by covalently linking the dirhodium core to an engineered prolyl oligopeptidase scaffold via strain-promoted azide-alkyne cycloaddition.[61] The newly formed ArM catalyzed the intermolecular cyclopropanation with excellent yield and enantioselectivity. (Figure 6a) This is the first example of dirhodium catalyzed asymmetric cyclopropanation using a protein scaffold as chiral environment. Later on, the same group achieved the directed evolution of the dirhodium ArM cyclopropanase via random mutagenesis.[62] Error prone PCR and combinatorial codon mutagenesis helped to highlight the importance of non-active site mutations for ArM optimization. The highly engineered POP scaffold with 12 mutations achieved up to 76 TON and 92% ee for the cyclopropanation reaction. (Figure 6b) It also showed improved enantioselectivity for other carbene insertion reactions such as Si-H, S-H and N-H insertions. This work highlighted the significant impact of distal mutations far away from active site to catalysis, highlighting the importance of random mutagenesis for ArM evolution.

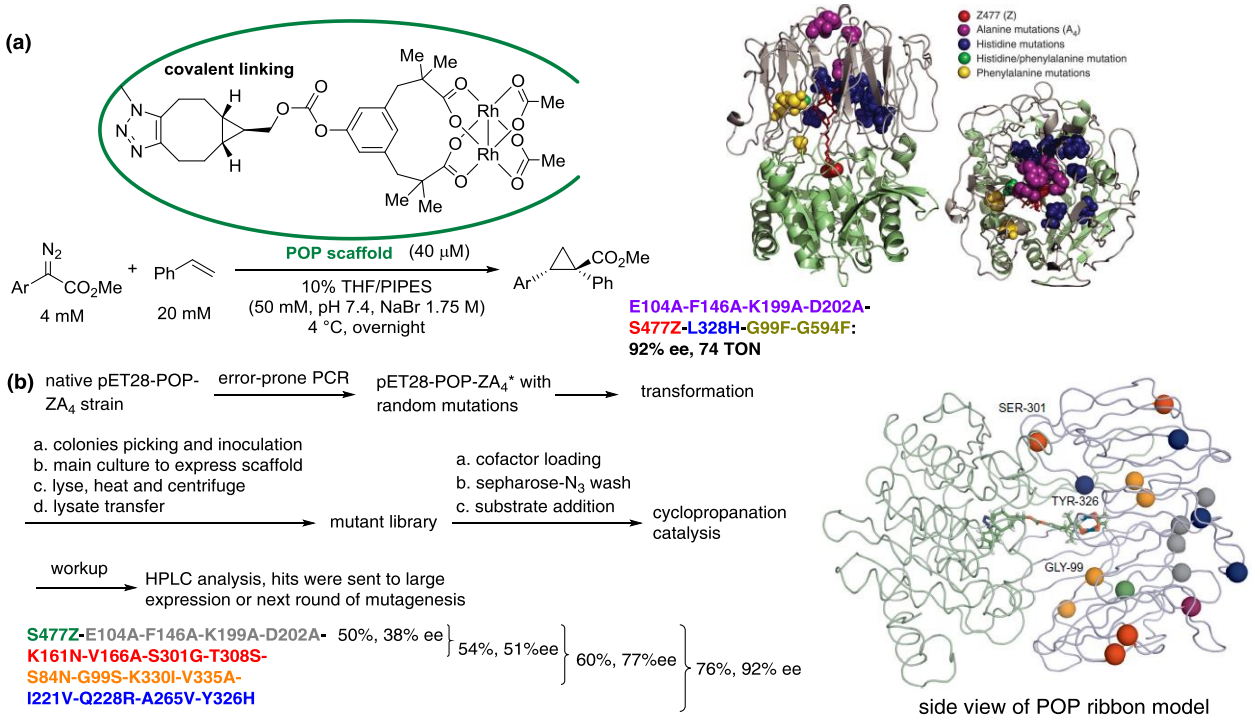


Figure 6. (a) Initial engineered dirhodium-POP ArM catalyzed asymmetric cyclopropanation. (b) Workflow of directed evolution of dirhodium cyclopropanase by random mutagenesis and selected evolution results. Mutations are highlighted in different colors: O-ZA_4 (grey), 1-NAGS (red), 2-NSIA (orange) and 3-VRVH (blue). S301, Y326 and G99 showed great impact on enantioselectivity for cyclopropanation.

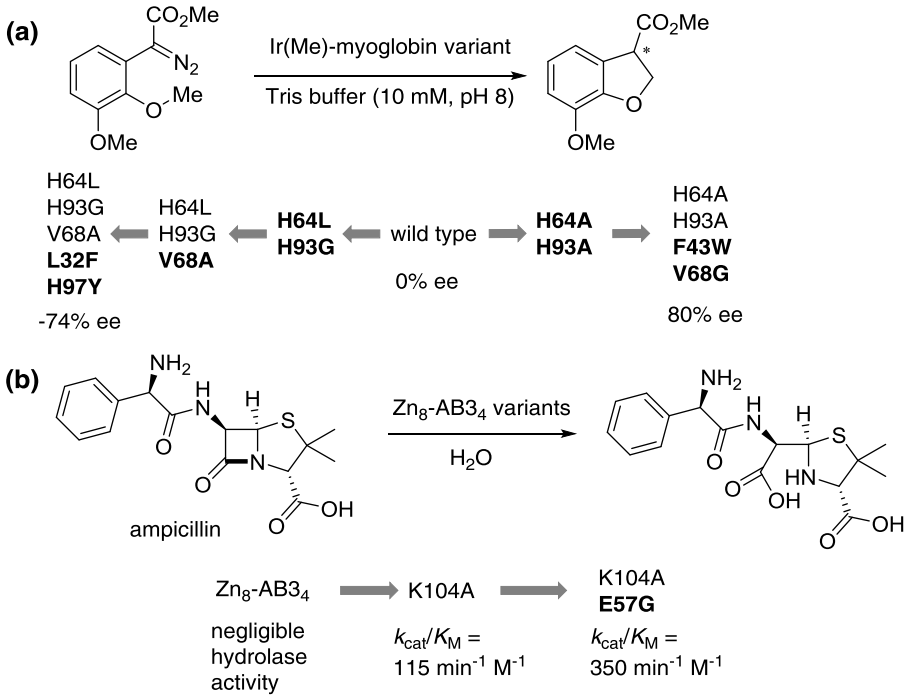
1.3.3 Summary

Dirhodium complexes are efficient and selective catalysts for carbene transfer reactions. Different substituted carboxylate ligands result in site selectivity and enantioselectivity. Introducing a dirhodium core into biomolecules to form complex biocompatible catalysts for asymmetric catalysis is a rather new and promising research area. Unlike the traditional chirality induction by bulky chiral carboxylate ligands coordinated to dirhodium moiety, it is difficult to predict how amino acid residues influence the dirhodium activity and selectivity even with an X-ray structure.

1.4 Directed evolution of ArMs

Directed evolution is a powerful method to optimize the structure and the function of enzymes. Research experts, such as Reetz[63], Arnold[64], Turner[65], Bornscheuer[66], Liu[67], Fasan[68] and etc, have reported outstanding work on the directed evolution of natural enzymes. However, it is difficult to perform directed evolution for ArMs using the traditional high-throughput screening methods. Because the formation of ArMs requires *in vitro* cofactor incorporation and validated screening assays as well as fast analysis methods need to be developed. Only in recent years, a few the directed evolution campaigns with ArMs have been reported.[17,62,69–72]

The Hartwig group achieved a non-natural C-H insertion by replacing the native Fe center of the myoglobin scaffold with Ir. Through engineering the apoprotein scaffold, the reconstituted Ir(Me)-myoglobin ArM showed significant selectivity improvement. (Scheme 12a) Tezcan and coworkers modified cytochrome cb_{562} to self-assemble through hydrophobic interactions into a tetrameric complex.[73] Introduction of a single mutation leads to a remarkable improvement of the hydrolysis activity toward *p*-nitrophenyl acetate and the ampicillin. (Scheme 12b) This hydrolase activity with ampicillin directly affected the cell survival on ampicillin-containing media, allowing the authors to use cell survival to improve the hydrolysis efficiency by directed evolution.



Scheme 12. (a) Evolutionary trajectory for diversifying and improving enantioselectivity of C-H insertion reaction. (b) Progression of evolution for hydrolyzing ampicillin.

The Ward group reported the directed evolution of artificial metathesase based on the biotin-streptavidin technology.[74] The *N*-terminal OmpA signal sequence allowed to secrete the expressed streptavidin into the periplasm, thus enabling the assembly of ArMs in the presence of whole *E. coli*

cells. This greatly accelerated the screening throughput. The umbelliferone precursor was selected as model substrate for the perplasmic screening and lead to the identification of a quintuple mutant V47A-N49K-T114Q-A119G-K121R. This mutant was purified and tested *in vitro* with other substrates for ring closing metathesis. Results showed moderate improvement in TON compared with WT.

In the process of the directed evolution of ArMs, some challenging issues still remain, such as synthetic cofactor deactivation by cellular components, expression level difference between mutant, and effective mutant library construction. There is always no guarantee that the tested reaction activity or selectivity could obtain significant improvement, so other techniques such as X-ray structure or computational predictions are needed to guide how the directed evolution proceeds.

1.5 Aim of the thesis

The aim of this thesis is to develop two new artificial metalloenzymes which incorporate transition metal Ru and Rh into two different protein scaffolds, and to engineer a previously studied artificial transfer hydrogenase which incorporates Ir within streptavidin by directed evolution, employing a self-immolative substrate for high throughput screening.

Chapter 2 outlines the efforts toward the development of an efficient artificial metathesase employing human carbonic anhydrase II as the scaffold. As hCA II has a deep cavity and binds to sulfonamide inhibitors with high affinity, various aryl-sulfonamide anchored ruthenium complexes were synthesized and tested in combination with hCA II mutants on different diolefin substrates. Ring closing metathesis served as the standard reaction to identify the best combinations of substrate, ArM and reaction buffer.

The study described in chapter 3 aims at creating efficient and selective dirhodium bound artificial carbenoid transferases based on biotin-streptavidin technology. By site-directed mutagenesis of the Sav scaffold, the influence of amino acid residues at specific positions for dirhodium catalyzed carbene transfer reactions was explored. As the dirhodium moiety proved to be stable and active in a cellular environment, the activity of the best dirhodium ArM was tested employing Sav-expressing *E. coli* cells for future high throughput screening purposes.

In chapter 4, the feasibility of improved activity based on an inserted loop around the biotin-binding vestibule of an artificial transfer hydrogenase is discussed. A streamlined screening platform was set up for fast evaluation of the studied constructs.

1.6 Reference

(1) Schwizer, F.; Okamoto, Y.; Heinisch, T.; Gu, Y.; Pellizzoni, M. M.; Lebrun, V.; Reuter, R.; Köhler, V.; Lewis, J. C.; Ward, T. R. Artificial Metalloenzymes: Reaction Scope and Optimization Strategies. *Chem. Rev.* **2018**, *118* (1), 142–231.

(2) Wieszczycka, K.; Staszak, K. Artificial Metalloenzymes as Catalysts in Non-Natural Compounds Synthesis. *Coord. Chem. Rev.* **2017**, *351*, 160–171.

(3) Yu, F.; Cangelosi, V. M.; Zastrow, M. L.; Tegoni, M.; Plegaria, J. S.; Tebo, A. G.; Mocny, C. S.; Ruckthong, L.; Qayyum, H.; Pecoraro, V. L. Protein Design: Toward Functional Metalloenzymes. *Chem. Rev.* **2014**, *114* (7), 3495–3578.

(4) Grimm, A. R.; Sauer, D. F.; Polen, T.; Zhu, L.; Hayashi, T.; Okuda, J.; Schwaneberg, U. A Whole Cell *E. Coli* Display Platform for Artificial Metalloenzymes: Poly(phenylacetylene) Production with a Rhodium–Nitrobindin Metalloprotein. *ACS Catal.* **2018**, *8* (3), 2611–2614.

(5) Li, L.-L.; Yuan, H.; Liao, F.; He, B.; Gao, S.-Q.; Wen, G.-B.; Tan, X.; Lin, Y.-W. Rational Design of Artificial Dye-Decolorizing Peroxidases Using Myoglobin by Engineering Tyr/Trp in the Heme Center. *Dalt. Trans.* **2017**, *46* (34), 11230–11238.

(6) Wang, B.; Bols, M. Artificial Metallooxidases from Cyclodextrin Diacids. *Chem. - Eur. J.* **2017**, *23* (55), 13766–13775.

(7) Hestericová, M.; Heinisch, T.; Alonso-Cotchico, L.; Maréchal, J. D.; Vidossich, P.; Ward, T. R. Directed Evolution of an Artificial Imine Reductase. *Angew. Chemie., Int. Ed.* **2018**, *57* (7), 1863–1868.

(8) Sreenilayam, G.; Moore, E. J.; Steck, V.; Fasan, R. Metal Substitution Modulates the Reactivity and Extends the Reaction Scope of Myoglobin Carbene Transfer Catalysts. *Adv. Synth. Catal.* **2017**, *359* (12), 2076–2089.

(9) Heinisch, T.; Ward, T. R. Artificial Metalloenzymes Based on the Biotin-Streptavidin Technology: Challenges and Opportunities. *Acc. Chem. Res.* **2016**, *49* (9), 1711–1721.

(10) Pàmies, O.; Diéguez, M.; Bäckvall, J.-E. Artificial Metalloenzymes in Asymmetric Catalysis: Key Developments and Future Directions. *Adv. Synth. Catal.* **2015**, 357 (8), 1567–1586.

(11) Yu, Y.; Hu, C.; Xia, L.; Wang, J. Artificial Metalloenzyme Design with Unnatural Amino Acids and Non-Native Cofactors. *ACS Catal.* **2018**, 8 (3), 1851–1863.

(12) Hu, C.; Yu, Y.; Wang, J. Improving Artificial Metalloenzymes' Activity by Optimizing Electron Transfer. *Chem. Commun.* **2017**, *53* (30), 4173–4186.

(13) Jeschek, M.; Panke, S.; Ward, T. R. Artificial Metalloenzymes on the Verge of New-to-Nature Metabolism. *Trends Biotechnol.* **2018**, *36* (1), 60–72.

(14) Upp, D. M.; Lewis, J. C. Selective C–H Bond Functionalization Using Repurposed or Artificial Metalloenzymes. *Curr. Opin. Chem. Biol.* **2017**, *37*, 48–55.

(15) Lin, Y. W. Rational Design of Metalloenzymes: From Single to Multiple Active Sites. *Coord.*

Chem. Rev. **2017**, *336*, 1–27.

(16) Oohora, K.; Meichin, H.; Kihira, Y.; Sugimoto, H.; Shiro, Y.; Hayashi, T. Manganese(V) Porphycene Complex Responsible for Inert C–H Bond Hydroxylation in a Myoglobin Matrix. *J. Am. Chem. Soc.* **2017**, *139* (51), 18460–18463.

(17) Key, H. M.; Dydio, P.; Clark, D. S.; Hartwig, J. F. Abiological Catalysis by Artificial Haem Proteins Containing Noble Metals in Place of Iron. *Nature* **2016**, *534* (7608), 534–537.

(18) Dydio, P.; Key, H. M.; Nazarenko, A.; Rha, J. Y.-E.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. An Artificial Metalloenzyme with the Kinetics of Native Enzymes. *Science* **2016**, *354* (6308), 102–106.

(19) Sreenilayam, G.; Moore, E. J.; Steck, V.; Fasan, R. Stereoselective Olefin Cyclopropanation under Aerobic Conditions with an Artificial Enzyme Incorporating an Iron-Chlorin e6 Cofactor. *ACS Catal.* **2017**, *7* (11), 7629–7633.

(20) Lupi, F.; Marzo, T.; D'Adamio, G.; Cretella, S.; Cardona, F.; Messori, L.; Goti, A. Diruthenium Diacetate Catalysed Aerobic Oxidation of Hydroxylamines and Improved Chemoselectivity by Immobilisation to Lysozyme. *ChemCatChem* **2017**, *9* (22), 4225–4230.

(21) Jarvis, A. G.; Obrecht, L.; Deuss, P. J.; Laan, W.; Gibson, E. K.; Wells, P. P.; Kamer, P. C. J. Enzyme Activity by Design: An Artificial Rhodium Hydroformylase for Linear Aldehydes. *Angew. Chemie., Int. Ed.* **2017**, *56* (44), 13596–13600.

(22) Drienovská, I.; Alonso-Cotchico, L.; Vidossich, P.; Lledós, A.; Maréchal, J.-D.; Roelfes, G. Design of an Enantioselective Artificial Metallo-Hydratase Enzyme Containing an Unnatural Metal-Binding Amino Acid. *Chem. Sci.* **2017**, *8* (10), 7228–7235.

(23) Bersellini, M.; Roelfes, G. Multidrug Resistance Regulators (MDRs) as Scaffolds for the Design of Artificial Metalloenzymes. *Org. Biomol. Chem.* **2017**, *15* (14), 3069–3073.

(24) Dey, S.; Rühl, C. L.; Jäschke, A. Catalysis of Michael Additions by Covalently Modified G-Quadruplex DNA. *Chem. - Eur. J.* **2017**, *23* (50), 12162–12170.

(25) Fujieda, N.; Nakano, T.; Taniguchi, Y.; Ichihashi, H.; Sugimoto, H.; Morimoto, Y.; Nishikawa, Y.; Kurisu, G.; Itoh, S. A Well-Defined Osmium-Cupin Complex: Hyperstable Artificial Osmium Peroxygenase. *J. Am. Chem. Soc.* **2017**, *139* (14), 5149–5155.

(26) Christianson, D. W. Carbonic Anhydrase: Evolution of the Zinc Binding Site by Nature and by Desian. *Acc. Chem. Res.* **1996**, *29* (7), 331–339.

(27) Lopez, M.; Vu, H.; Wang, C. K.; Wolf, M. G.; Groenhof, G.; Innocenti, A.; Supuran, C. T.; Poulsen, S.-A. Promiscuity of Carbonic Anhydrase II. Unexpected Ester Hydrolysis of Carbohydrate-Based Sulfamate Inhibitors. *J. Am. Chem. Soc.* **2011**, *133* (45), 18452–18462.

(28) Martin, D. P.; Hann, Z. S.; Cohen, S. M. Metalloprotein-Inhibitor Binding: Human Carbonic Anhydrase II as a Model for Probing Metal-Ligand Interactions in a Metalloprotein Active Site. *Inorg. Chem.* **2013**, *52* (21), 12207–12215.

(29) Roy, B. C.; Banerjee, A. L.; Swanson, M.; Jia, X. G.; Haldar, M. K.; Mallik, S.; Srivastava, D.

K. Two-Prong Inhibitors for Human Carbonic Anhydrase II. *J. Am. Chem. Soc.* **2004**, *126* (41), 13206–13207.

(30) Sippel, K. H.; Robbins, A. H.; Domsic, J.; Genis, C.; Agbandje-Mckenna, M.; McKenna, R. High-Resolution Structure of Human Carbonic Anhydrase II Complexed with Acetazolamide Reveals Insights into Inhibitor Drug Design. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2009**, *65* (10), 992–995.

(31) Fernández-Gacio, A.; Codina, A.; Fastrez, J.; Riant, O.; Soumillion, P. Transforming Carbonic Anhydrase into Epoxide Synthase by Metal Exchange. *ChemBioChem* **2006**, *7* (7), 1013–1016.

(32) Jing, Q.; Okrasa, K.; Kazlauskas, R. J. Stereoselective Hydrogenation of Olefins Using Rhodium-Substituted Carbonic Anhydrase - A New Reductase. *Chem. - Eur. J.* **2009**, *15* (6), 1370–1376.

(33) Jing, Q.; Kazlauskas, R. J. Regioselective Hydroformylation of Styrene Using Rhodium-Substituted Carbonic Anhydrase. *ChemCatChem* **2010**, *2* (8), 953–957.

(34) Kim, M.-C.; Lee, S.-Y. Catalytic Water Oxidation by Iridium-Modified Carbonic Anhydrase. *Chem. - Asian J.* **2018**, *13* (3), 334–341.

(35) Key, H. M.; Clark, D. S.; Hartwig, J. F. Generation, Characterization, and Tunable Reactivity of Organometallic Fragments Bound to a Protein Ligand. *J. Am. Chem. Soc.* **2015**, *137* (25), 8261–8268.

(36) Monnard, F. W.; Nogueira, E. S.; Heinisch, T.; Schirmer, T.; Ward, T. R. Human Carbonic Anhydrase II as Host Protein for the Creation of Artificial Metalloenzymes: The Asymmetric Transfer Hydrogenation of Imines. *Chem. Sci.* **2013**, *4* (8), 3269.

(37) Heinisch, T.; Pellizzoni, M.; Dürrenberger, M.; Tinberg, C. E.; Köhler, V.; Klehr, J.; Häussinger, D.; Baker, D.; Ward, T. R. Improving the Catalytic Performance of an Artificial Metalloenzyme by Computational Design. *J. Am. Chem. Soc.* **2015**, *137* (32), 10414–10419.

(38) Kean, K. M.; Porter, J. J.; Mehl, R. A.; Karplus, P. A. Structural Insights into a Thermostable Variant of Human Carbonic Anhydrase II. *Protein Sci.* **2018**, *27* (2), 573–577.

(39) Supuran, C. T.; Scozzafava, A.; Casini, A. Carbonic Anhydrase Inhibitors. *Med. Res. Rev.* **2003**, *23* (2), 146–189.

(40) Supuran, C. T. Carbonic Anhydrases: Novel Therapeutic Applications for Inhibitors and Activators. *Nat. Rev. Drug Discov.* **2008**, *7* (2), 168–181.

(41) Krall, N.; Pretto, F.; Decurtins, W.; Bernardes, G. J. L.; Supuran, C. T.; Neri, D. A Small-Molecule Drug Conjugate for the Treatment of Carbonic Anhydrase IX Expressing Tumors. *Angew. Chem., Int. Ed.* **2014**, *53* (16), 4231–4235.

(42) Davies, H. M. L.; Morton, D. Guiding Principles for Site Selective and Stereoselective Intermolecular C–H Functionalization by Donor/acceptor Rhodium Carbenes. *Chem. Soc. Rev.* **2011**, *40* (4), 1857.

(43) Selander, N.; Worrell, B. T.; Chuprakov, S.; Velaparthi, S.; Fokin, V. V. Arylation of

rhodium(II) Azavinyl Carbenes with Boronic Acids. *J. Am. Chem. Soc.* **2012**, *134* (36), 14670–14673.

(44) Xu, X.; Zavalij, P. Y.; Doyle, M. P. Catalytic Asymmetric Syntheses of Quinolizidines by Dirhodium-Catalyzed Dearomatization of Isoquinolinium/pyridinium Methylides-the Role of Catalyst and Carbene Source. *J. Am. Chem. Soc.* **2013**, *135* (33), 12439–12447.

(45) Tong, L. H.; Clifford, S.; Gomila, A.; Duval, S.; Guénée, L.; Williams, A. F. Supramolecular Squares of Dirhodium(ii) Tetracarboxylate: Combining Carboxylate-Exchange and Metal–ligand Coordination for Self-Assembly. *Chem. Commun.* **2012**, *48* (79), 9891.

(46) Kumar, D. K.; Filatov, A. S.; Napier, M.; Sun, J.; Dikarev, E. V.; Petrukhina, M. A. Dirhodium Paddlewheel with Functionalized Carboxylate Bridges: New Building Block for Self-Assembly and Immobilization on Solid Support. *Inorg. Chem.* **2012**, *51* (8), 4855–4861.

(47) Wong, F. M.; Wang, J.; Hengge, A. C.; Wu, W. Mechanism of Rhodium-Catalyzed Carbene Formation from Diazo Compounds. *Org. Lett.* **2007**, *9* (9), 1663–1665.

(48) Werlé, C.; Goddard, R.; Philipps, P.; Farès, C.; Fürstner, A. Structures of Reactive Donor/Acceptor and Donor/Donor Rhodium Carbenes in the Solid State and Their Implications for Catalysis. *J. Am. Chem. Soc.* **2016**, *138* (11), 3797–3805.

(49) Davies, H. M. L.; Lian, Y. The Combined C-H Functionalization/cope Rearrangement: Discovery and Applications in Organic Synthesis. *Acc. Chem. Res.* **2012**, *45* (6), 923–935.

(50) Liao, K.; Pickel, T. C.; Boyarskikh, V.; Bacsa, J.; Musaev, D. G.; Davies, H. M. L. Site-Selective and Stereoselective Functionalization of Non-Activated Tertiary C-H Bonds. *Nature* **2017**, *551* (7682), 609–613.

(51) Liao, K.; Negretti, S.; Musaev, D. G.; Bacsa, J.; Davies, H, M. L. Site-Selective and Stereoselective Functionalization of Unactivated C–H Bonds. *Nature* **2016**, *533* (7602), 230–234.

(52) Wynne, D. C.; Olmstead, M. M.; Jessop, P. G. Supercritical and Liquid Solvent Effects on the Enantioselectivity of Asymmetric Cyclopropanation with Tetrakis [1- [(4-Tert-Butylphenyl) - Sulfonyl] - (2S) -Pyrrolidinecarboxylate] Dirhodium (II). *J. Am. Chem. Soc.* **2000**, *122* (32), 7638–7647.

(53) Davies, H. M..; Antoulinakis, E. G. Recent Progress in Asymmetric Intermolecular C–H Activation by Rhodium Carbenoid Intermediates. *J. Organomet. Chem.* **2001**, 617–618, 47–55.

(54) Davies, H. M. L.; Morton, D. Guiding Principles for Site Selective and Stereoselective Intermolecular C-H Functionalization by Donor/acceptor Rhodium Carbenes. *Chem. Soc. Rev.* **2011**, *40* (4), 1857–1869.

(55) Popp, B. V.; Ball, Z. T. Structure-Selective Modification of Aromatic Side Chains with Dirhodium Metallopeptide Catalysts. *J. Am. Chem. Soc.* **2010**, *132* (19), 6660–6662.

(56) Sambasivan, R.; Ball, Z. T. Metallopeptides for Asymmetric Dirhodium Catalysis. *J. Am. Chem. Soc.* **2010**, *132* (27), 9289–9291.

(57) Sambasivan, R.; Ball, Z. T. Screening Rhodium Metallopeptide Libraries "on Bead": Asymmetric Cyclopropanation and a Solution to the Enantiomer Problem. *Angew. Chemie., Int. Ed.* **2012**, *51* (34), 8568–8572.

(58) Chen, Z.; Popp, B. V.; Bovet, C. L.; Ball, Z. T. Site-Specific Protein Modification with a Dirhodium Metallopeptide Catalyst. *ACS Chem. Biol.* **2011**, *6* (9), 920–925.

(59) Tishinov, K.; Schmidt, K.; Häussinger, D.; Gillingham, D. G. Structure-Selective Catalytic Alkylation of DNA and RNA. *Angew. Chemie., Int. Ed.* **2012**, *51* (48), 12000–12004.

(60) Bachmann, D. G.; Schmidt, P. J.; Geigle, S. N.; Chougnet, A.; Woggon, W. D.; Gillingham, D. G. Modular Ligands for Dirhodium Complexes Facilitate Catalyst Customization. *Adv. Synth. Catal.* **2015**, *357* (9), 2033–2038.

(61) Srivastava, P.; Yang, H.; Ellis-Guardiola, K.; Lewis, J. C. Engineering a Dirhodium Artificial Metalloenzyme for Selective Olefin Cyclopropanation. *Nat. Commun.* **2015**, *6*, 1–8.

(62) Yang, H.; Swartz, A. M.; Park, H. J.; Srivastava, P.; Ellis-Guardiola, K.; Upp, D. M.; Lee, G.; Belsare, K.; Gu, Y.; Zhang, C.; Mollering, R.; Lewis, J. C. Evolving Artificial Metalloenzymes via Random Mutagenesis. *Nat. Chem.* **2018**, *10* (3), 318–324.

(63) Wang, J.; Li, G.; Reetz, M. T. Enzymatic Site-Selectivity Enabled by Structure-Guided Directed Evolution. *Chem. Commun.* **2017**, *53* (28), 3916–3928.

(64) Chen, K.; Arnold, F. H. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc. Natl. Acad. Sci.* **1993**, *90* (12), 5618–5622.

(65) Turner, N. J. Directed Evolution Drives the next Generation of Biocatalysts. *Nat. Chem. Biol.* **2009**, *5* (8), 567–573.

(66) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the Third Wave of Biocatalysis. *Nature* **2012**, *485* (7397), 185–194.

(67) Packer, M. S.; Liu, D. R. Methods for the Directed Evolution of Proteins. *Nat. Rev. Genet.* **2015**, *16* (7), 379–394.

(68) Bordeaux, M.; Tyagi, V.; Fasan, R. Highly Diastereoselective and Enantioselective Olefin Cyclopropanation Using Engineered Myoglobin-Based Catalysts. *Angew. Chemie., Int. Ed.* **2015**, *54* (6), 1744–1748.

(69) Ilie, A.; Reetz, M. T. Directed Evolution of Artificial Metalloenzymes. *Isr. J. Chem.* **2015**, *55* (1), 51–60.

(70) Ward, T. R. Directed Evolution of Iridium-Substituted Myoglobin Affords Versatile Artificial Metalloenzymes for Enantioselective C−C Bond-Forming Reactions. *Angew. Chemie., Int. Ed.* **2016**, *55* (48), 14909–14911.

(71) Hyster, T. K.; Ward, T. R. Genetic Optimization of Metalloenzymes: Enhancing Enzymes for Non-Natural Reactions. *Angew. Chemie., Int. Ed.* **2016**, *55* (26), 7344–7357.

(72) Obexer, R.; Godina, A.; Garrabou, X.; Mittl, P. R. E.; Baker, D.; Griffiths, A. D.; Hilvert, D.

Emergence of a Catalytic Tetrad during Evolution of a Highly Active Artificial Aldolase. *Nat. Chem.* **2017**, *9* (1), 50–56.

(73) Song, W. J.; Tezcan, F. A. A Designed Supramolecular Protein Assembly with *in Vivo* Enzymatic Activity. *Science* **2014**, *346* (6216), 1525–1528.

(74) Jeschek, M.; Reuter, R.; Heinisch, T.; Trindler, C.; Klehr, J.; Panke, S.; Ward, T. R. Directed Evolution of Artificial Metalloenzymes for in Vivo Metathesis. *Nature* **2016**, *537* (7622), 661–665.

Chapter 2 | Carbonic anhydrase II as host protein for the creation of a biocompatible artificial metathesase

Jingming Zhao, Anna Kajetanowicz and Thomas R. Ward*

Department of Chemistry, University of Basel, Spitalstrasse 51, CH-4056 Basel, Switzerland.

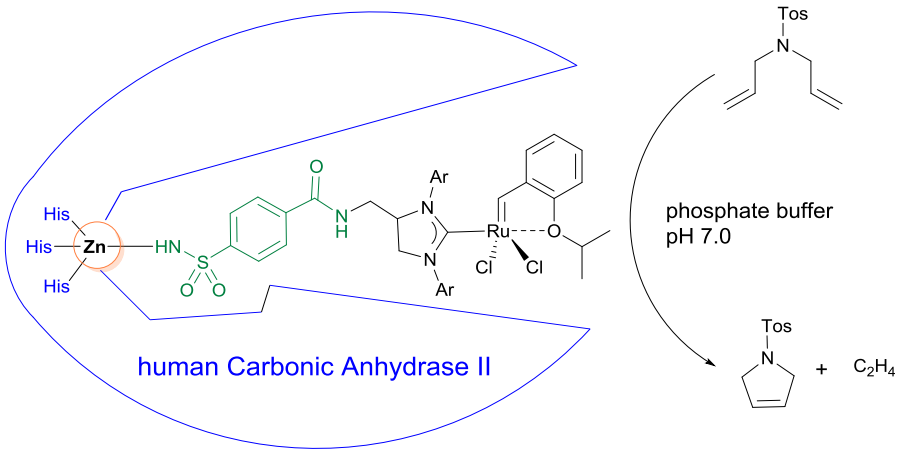
This work was published in *Org. Biomol. Chem.*, **2015**, *13*, 5662-5665.

2.1 Abstract

An artificial metathesase results from incorporation of an Hoveyda-Grubbs catalyst bearing an arylsulfonamide anchor within human carbonic anhydrase II. The optimization of the catalytic performance is achieved upon combining both chemical and genetic means. Up to 28 TONs were obtained within four hours under aerobic physiological conditions.

2.2 Introduction

Artificial metalloenzymes result from the incorporation of an abiotic cofactor within a host protein.[1] With biomedical applications in mind, it would be desirable to capitalize on a host protein which is overexpressed on the surface of cancer cells. Accumulation of the abiotic cofactor, which displays high affinity for the latter protein, may allow to site-specifically uncage a drug.[2] In this context, the ring-closing metathesis (RCM) is an attractive reaction as unactivated diolefins can be viewed as bioorthogonal. Furthermore, the intramolecular nature of the RCM may facilitate the reaction under highly dilute aqueous conditions.[3] Herein, we report on our efforts to exploit human Carbonic Anhydrase II (hCA II hereafter) for the creation of a biocompatible artificial metathesase, Scheme 1. Certain forms of cancer overexpress hCA IX, a membrane bound variant of hCA. These arylsulfonamide binding proteins are thus privileged targets for cancer therapy.[2,4a]

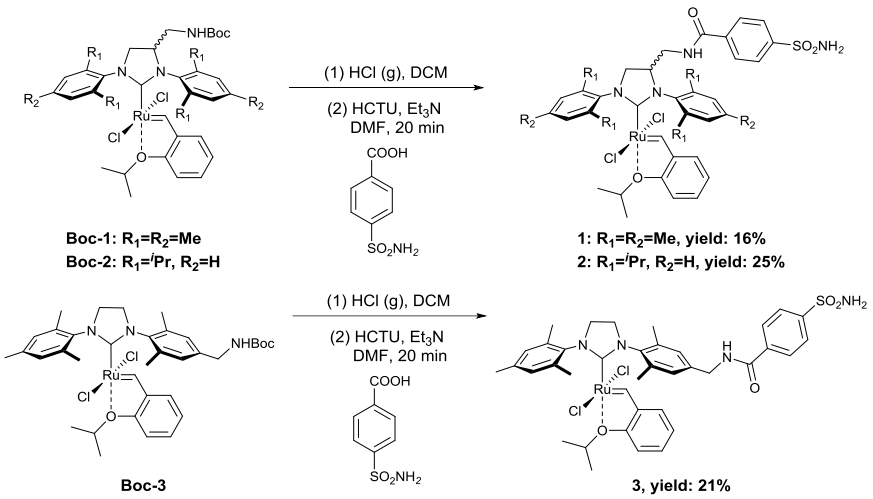


Scheme 1. Artificial metalloenzyme for ring-closing metathesis. Tethering an arylsulfonamide anchor (green) to an Hoveyda-Grubbs type catalyst (black) ensures the localization of the metal moiety within human Carbonic Anhydrase II (blue).

2.3 Results and discussion

Introduction of an arylsulfonamide-anchor on an Hoveyda-Grubbs 2nd generation-type catalyst ensures its localization within carbonic anhydrase.[4] For this purpose, complexes **Boc-1**, **Boc-2** and **Boc-3**,[5]

were deprotected *in situ* and reacted with 4-sulfamoylbenzoic acid to afford the corresponding sulfonamide-bearing metathesis cofactors **1**, **2** and **3**, Scheme 2 (see ESI† for full Experimental details).



Scheme 2. Synthesis of olefin metathesase cofactors **1**-**3** bearing an arylsulfonamide anchor for incorporation within hCA II.

The catalytic performance of the artificial metathesases was evaluated using the ring-closing metathesis of *N*-tosyl diallyl-amine in the presence of 1 mol% ruthenium. To ensure a homogeneous mixture, water/DMSO (9/1) was selected.

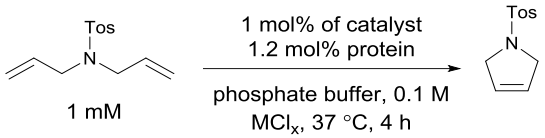
Comparison of catalysts **1**–**3** in the absence of hCA II at pH 6.0 in the presence of 0.1 M $MgCl_2$ reveals that the bulkiest catalyst **2** outperforms catalysts **1** and **3** (Table 1, entries 1–3). The same trend is observed upon incorporation of the cofactors **1**–**3** into WT hCA II (Table 1, entries 4–6). Catalyst **2** was thus selected for further optimization. With no $MgCl_2$ added and at pH 7.0, catalyst **2** and the corresponding metathesase **2** ⊂ WT hCA II afforded 23 and 14 turnovers after four hours at 37 °C. Performing catalysis under strict exclusion of oxygen yielded very similar results.

Reactions carried out at pH 7.0 and in the presence of 154 mM NaCl (corresponding to physiological conditions) yielded 32 and 21 TONs for **2** and **2** ⊂ WT hCA II respectively (Table 1, entries 15 and 16). As can be appreciated, the TON of the catalyst is pH dependent, both in the presence and in the absence of hCA II. The best performance is obtained at lower pH and high salt concentration (Table 1, entries 12 and 13). As for other artificial metalloenzymes, we do not believe that the pI of the host rotein influences significantly the catalyst performance.

Compared to the other four artificial metathesases reported to date,[3a–e] the system presented herein presents the following advantageous features (Table 2): (i) it does not require an inert atmosphere; (ii) the substrate concentration is the lowest of all systems reported to date; (iii) except for the metathesase based on FhuA (which requires SDS, a surfactant), it displays the highest turnover frequency and (iv) it catalyzes RCM at pH 7.0, temperature 37 °C and physiological [NaCl] concentrations. These results

thus suggest that WT hCA II is a suitable host for the creation of artificial metathesases operating under physiological conditions and at low catalyst concentrations (i.e.10 μM).

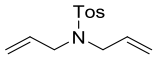
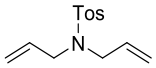
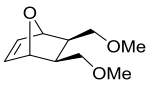
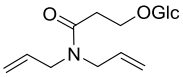
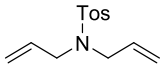
Table 1 Selected results for the ring-closing metathesis of *N*-tosyl diallylamine.[a]



Entry	Catalyst	HCA II	pH	[MCl_x] mol/L	TON[b]
1	**1**	–	6.0	$MgCl_2$ 0.1	20 ± 0.4
2	**2**	–	6.0	$MgCl_2$ 0.1	48 ± 0.8
3	**3**	–	6.0	$MgCl_2$ 0.1	25 ± 0.7
4	**1**	WT	6.0	$MgCl_2$ 0.1	13 ± 1.3
5	**2**	WT	6.0	$MgCl_2$ 0.1	45 ± 2.0
6	**3**	WT	6.0	$MgCl_2$ 0.1	16 ± 1.0
7	**2**	WT	6.0	$MgCl_2$ 0.1	40 ± 1.5
8	**2**	WT	7.0	$MgCl_2$ 0.1	28 ± 1.1
9	**2**	–	7.0	–	23 ± 2.1
10	**2**	WT	7.0	–	14 ± 0.5
11[c]	**2**	WT	7.0	–	20 ± 2.3
12	**2**	–	5.0	$MgCl_2$ 0.5	85 ± 1.0
13	**2**	WT	5.0	$MgCl_2$ 0.5	78 ± 2.5
14	**2**	WT	6.0	–	23 ± 2.6
15	**2**	WT	8.0	–	21 ± 1.2
16	**2**	–	7.0	NaCl 0.154	32 ± 2.0
17	**2**	WT	7.0	NaCl 0.154	21 ± 1.8
18	**2**	WT	7.0	NaCl 0.5	32 ± 1.8
19	**2**	WT	7.0	NaCl 1.0	29 ± 1.2
20	**2**	I91A	7.0	–	18 ± 3.3
21	**2**	F131A	7.0	–	16 ± 1.3
22	**2**	L198F	7.0	–	18 ± 1.6
23	**2**	L198H	7.0	–	22 ± 0.1
24	**2**	L198H	7.0	NaCl 0.154	28 ± 0.6
25	**2**	L198A	7.0	–	15 ± 1.7
26	**2**	L198Q	7.0	–	14 ± 0.1
27	**2**	K170A	7.0	–	15 ± 2.0

[a] Reaction conditions: [substrate] = 1 mM, [catalyst] = 10 μM, [hCA II] = 12 μM, V_{tot} 200 μL (V_{DMSO} 20 μL), 37 °C for 4 hours. The reactions were carried out in triplicate. Very similar results were obtained under rigorous exclusion of oxygen. [b] Turnover number. [c] [substrate] = 5 mM, [catalyst] = 50 μM, [hCA II] = 60 μM.

Table 2 Summary of the catalytic performance of artificial metathesases reported to date

	Hilvert[3b]	Ward[3a]	Schwaneberg[3d, 3e]	Matsuo[3c]	Ward
[substrate]	 5 mM	 15.21 mM	 100 mM	 8 mM	 1 mM
Reaction type	RCM	RCM	ROMP	RCM	RCM
Anchoring of Ru-cofactor	covalent	supramolecular	covalent	covalent	dative
Host protein	MjHSP[1] 4 mol%	Avidin 4.8 mol%	FhuA ΔDCVF^{tev}[2] 0.08 mol%	α-Chymotrypsin 0.63 mol%	hCA II 1 mol%
Temp. Time pH	45 °C 12 h 2	40 °C 16 h 4	25 °C 68 h 7	25 °C 2 h 7	37 °C 4 h 7
Reaction conditions	10 mM HCl, Water/*t*-BuOH 4/1 under air	0.1 M acetate Water/DMSO 5/1 0.5 M $MgCl_2$, under air	Water/THF 9/1 SDS 1% under N_2	Degassed 100 mM KCl, under N_2	0.1 M phosphate Water/DMSO 9/1 under air
TON	25	20	955	20	28

[1] MjHSP: *M. jannaschii* small heat shock protein.
[2] FhuA ΔDCVF^{tev}: engineered variants of the β-barrel ferric hydroxamate uptake protein component A

To gain insight into the localization of *rac*-**2** within WT hCA II, both enantiomers were docked using the GOLD program,[7] Fig. 1. As can be appreciated, the cofactor fits nicely within the cone-shaped binding funnel of hCA II, presenting its alkyl-idene moiety at the surface of the protein (Fig. 1). With the aim of improving the TON of the artificial metathesase, residues I91, F131, L198, K170 were subjected to site-directed mutagenesis (Table 1, entries 20–27). Lypophylic, polar and potentially coordinating aminoacid residues were engineered into these positions. A selection of mutants tested is presented in Table 1, entries 20–27. To our delight, **2** ⊂ hCA II L198H yielded significantly improved catalytic performance: up to 28 TONs under physiological conditions (Table 1, entry 24). Considering that both His198 and Phe198 afford similar TONs, we do not believe that the former residue coordinates to ruthenium. At this point however, it is difficult to rationalize or predict the effect a point mutations on the outcome of catalysis.

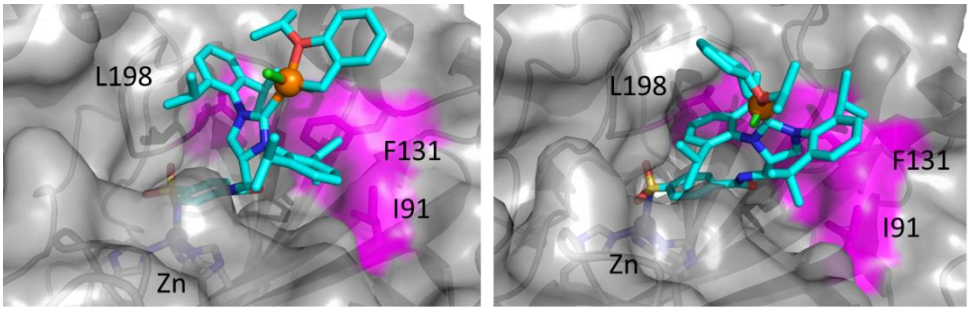


Figure 1. Docked structure of (*R*)-**2** ⊂ WT hCA II (a) and (*S*)-**2** ⊂ WT hCA II (b). The ruthenium cofactor is displayed as stick and hCA II as solvent accessible transparent surface. Residues subjected to mutagenesis are highlighted in magenta.

In order to ensure localisation of the cofactor **2** within hCA II under catalytic conditions, its affinity was determined using the dansylamide displacement assay.[8] Dansylamide (DNSA) displays enhanced fluorescence upon incorporation within hCA II. The non-covalent probe can be displaced by high affinity arylsulfonamide-bearing hCA II inhibitors, leading to a decrease in fluorescence.

For DNSA, we obtained a K_d = 4.83 µM and 17.35 µM for WT hCA II and L198H hCA II respectively. The displacement assay (see ESI† for details) yielded K_d = 90.40 nM and 205.10 nM for **2** ⊂ WT hCA II and **2** ⊂ hCA II L198H respectively, Fig. 2. Although we cannot exclude additional non-specific binding on the surface of hCA II, we feel that the exquisite specificity of arylsulfonamides most probably ensures selective binding to the $Zn(His)_3$ moiety. We tentatively attribute the modest quality of the fit to the use of a *rac*-**2** (and thus the presence of two inhibitors with potentially different affinities). Under catalytic conditions 96% and 93% of **2** is bound to WT hCA II and hCA II L198H respectively.

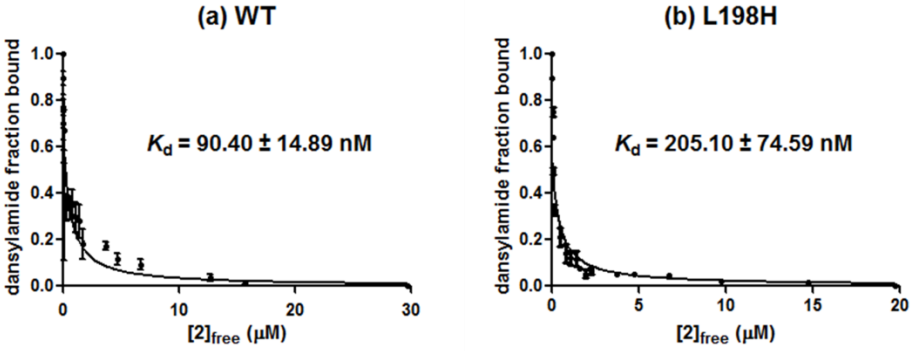


Figure 2. Determination of the dissociation constant of **2** ⊂ hCA II using the dansylamide displacement assay: WT hCA II (a) and L198H hCA II (b) [hCA II] = 0.25 µM, [Dansylamide] = 20 µM, See SI for full experimental details. All measurements were performed in duplicate.

2.4 Conclusion

In summary, we have developed an artificial metathesase relying on hCA II as host protein. Importantly, the present system operates under aerobic physiological conditions and at low catalyst concentrations (i.e.10 µM). Current efforts are directed at evaluating hCA IX, a cell-surface hCA variant which is overexpressed in various forms of cancer.

2.5 Supporting information

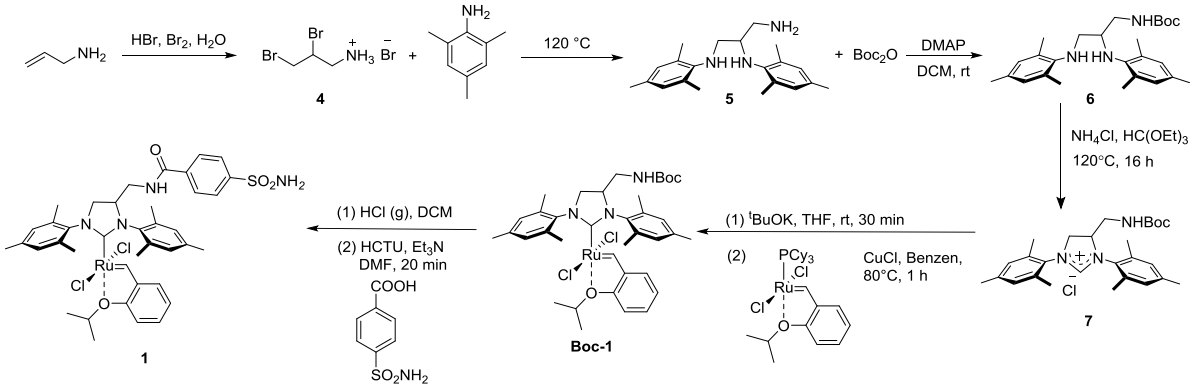
General aspects

Materials and reagents were purchased from the highest commercially available grade and used without further purification.

^{1}H and ^{13}C spectra were recorded on a Bruker 400 MHz, 500 MHz and 600 MHz Chemical shifts are reported in ppm (parts per million). Signals are quoted as s (singlet), d (doublet), t (triplet), brs (broad) and m (multiplet). Electron Spray Ionization Mass Spectra (ESI-MS) were recorded on a Bruker FTMS 4.7T bioAPEX II. Analysis of the catalytic runs was performed on an Agilent 1100 reverse phase HPLC. High performance liquid chromatography was performed on Agilent 1100 Series with UV-Vis detection.

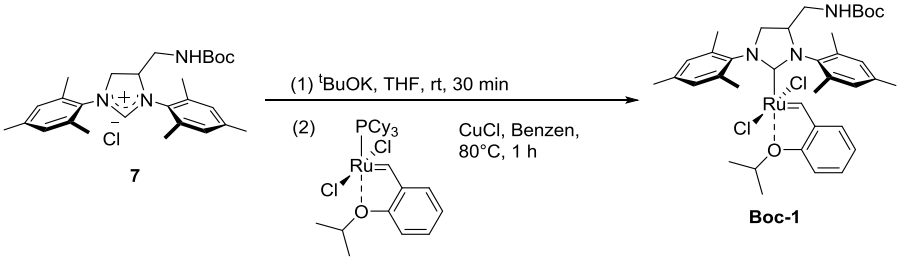
hCA II was expressed in 1 L shaking-flasks in an Infors HT Ecotron shaker and culture growth was monitored by UV-Vis at 600 nm with a Varian Cary 50 Scan. The protein was purified via sulphonamide affinity chromatography by using AKTA prime (Amersham Pharmacia Biotech, Software: PrimeView 5.0). The competitive displacement assays were performed using a Tecan Safire spectrophotometer using NUNC 96-well plates. The data were analyzed with Prism 5.0 software.

Catalyst synthesis



Compound **4**, **5**, **6**, **7**, **Boc-1** were prepared according to the reported procedures.[9]

Boc-1

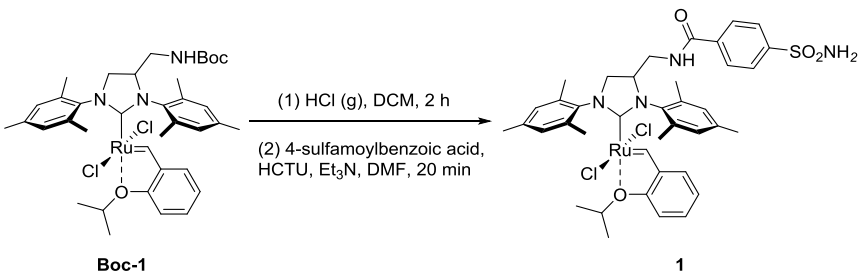


A suspension of ^{t}BuOK (112 mg, 0.99 mmol) was added to the imidazolinium salt **7** (471 mg, 0.99 mmol) in dry and degassed THF (15 ml). The mixture was stirred 30 min at room temperature. The yellow solution was transferred into a flask containing Hoveyda-Grubbs 1st generation (500 mg, 0.83 mmol) in benzene (75 ml) and stirred at 80 °C. After 1 hour, CuCl (99 mg, 0.99 mmol) was slowly added to the solution and stirred for an additional hour at 80 °C. The solvent was evaporated and purification was performed by chromatography (10% EtOAc in cyclohexane) to yield 221 mg (29%) of product **Boc-1** as a green powder.

^{1}H NMR (400 MHz, CD_2Cl_2): δ 16.33 (s, 1H), 7.49-7.45 (m, 1H), 7.01-6.98 (m, 4H), 6.87-6.80 (m, 2H), 6.76 (d, J = 8 Hz, 1H), 4.85-4.79 (m, 1H), 4.48-4.41 (m, 1H), 4.18 (t, J = 12 Hz, 1H), 3.86 (dd, J = 9.2 Hz, 6.8 Hz, 1H), 3.38-3.25 (m, 2H), 2.34 (s, 14 H), 1.35 (s, 4H), 1.30 (s, 9H), 1.18-1.16 (m, 6 H).

^{13}C NMR (100 Hz, CD_2Cl_2): δ 296.6, 214.5, 156.4, 152.4, 145.5, 139.4, 138.9, 130.5, 130.0, 129.7, 122.7, 122.6, 113.3, 79.9, 75.6, 64.3, 55.8, 43.2, 32.3, 30.1, 29.8, 28.4, 27.3, 23.1, 21.4, 21.4, 21.2, 21.2, 21.2, 18.5, 14.3.

Catalyst **1**



Boc-1 **1**

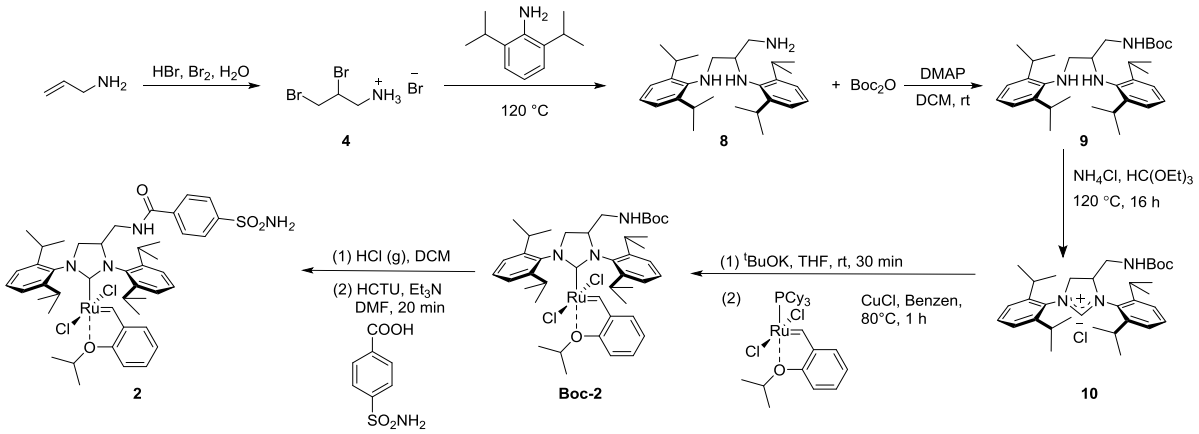
Boc-1 (45.3 mg, 0.06 mmol, 1 eq.) was dissolved in CH_2Cl_2 (2 mL). Gaseous HCl (H_2SO_4 added drop-wise to NH_4Cl) was bubbled through the solution and the mixture was stirred for 2 h at room temperature. The deprotection was monitored by thin-layer chromatography (TLC) (cyclohexane/ EtOAc 8:1). After consumption of the starting material, the solvent was evaporated under N_2 and the resulting mixture was dissolved in DMF (1 mL).
In another flask, 4-sulfamoylbenzoic acid (12.1 mg, 0.06 mmol, 1 eq.) and the coupling agent HCTU (27.3 mg, 0.066 mmol, 1.1 eq.) were vigorously stirred in 1 mL of DMF for 20 min. To this latter flask, the deprotected complex in DMF was added, followed by Et_3N (84.3 µL, 0.6 mmol, 10 eq.). The resulting mixture was stirred for 20 min at room temperature. The solvent was removed under vacuum. Purification by chromatography (50% EtOAc in cyclohexane) yielded 8 mg (16%) of catalyst **1** as a green powder.

^{1}H NMR (600 MHz, CD_2Cl_2): δ 16.37 (s, 1H), 7.39-7.88 (m, 4H), 7.58-7.55 (m, 1H), 7.14 (s, 1H), 7.09 (br, 2H), 6.96 (s, 1H), 6.92-6.91 (m, 2H), 6.87 (d, J = 8.6 Hz, 1H), 4.97-4.94 (m, 3H), 4.67 (br, 1H), 4.34 (t, J = 11.0 Hz, 1H), 3.95-3.92 (m, 1H), 3.63 (br, 1H), 2.88 (br, 2H), 2.46-2.40 (m, 14H), 1.43-1.27 (m, 10H).

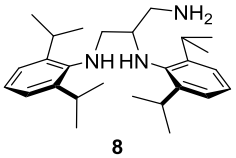
^{13}C NMR (150 Hz, CD_2Cl_2): δ 297.2, 211.0, 166.8, 152.5, 145.2, 137.8, 130.7, 130.4, 130.4, 130.0, 128.6, 126.8, 122.7, 113.3, 75.6, 64.0, 55.7, 55.7, 32.0, 29.9, 21.4, 21.1, 19.1.

HRMS (ESI, pos.) *m/z*: $[M-Cl]^+$ calcd for $C_{39}H_{46}ClN_4O_4RuS$, 803.1971; found: 803.1969.

Compound **8**, **9**, **10**, **Boc-2** were prepared using the same procedures as described for compound **5**, **6**, **7**, **Boc-1**.



***N*^{1}, *N*^{2}-bis(2,6-diisopropylphenyl)propane-1,2,3-triamine (8)**



8

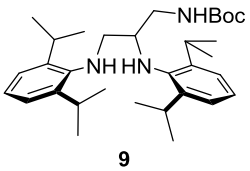
Compound **8** (yield 253 mg, 18 %) was prepared from the same procedure of compound **5**.[1]

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.02-6.88 (m, 6H), 3.38-3.04 (m, 9H), 3.01-2.93 (m, 3H), 2.85 (dd, J = 12.0 Hz, 5.2 Hz, 1H), 1.17-1.04 (m, 24H).

^{13}C NMR (100 Hz, $CDCl_3$): δ 142.7, 142.6, 141.9, 140.4, 124.0, 123.7, 123.6, 123.5, 59.9, 54.3, 53.4, 44.0, 27.9, 27.7, 24.3, 24.2, 24.1.

HRMS (ESI, pos.) *m/z*: $[M+H]^+$ calcd for $C_{27}H_{44}N_3$, 410.3536; found: 410.3529.

***tert*-butyl (2,3-bis((2,6-diisopropylphenyl)amino)propyl)carbamate (9)**



9

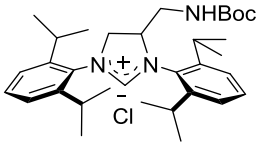
Compound **9** (yield 49 mg, 36 %) was prepared from the same procedure of compound **6**.

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.03-6.93(m, 6H), 4.96 (br, 1H), 3.45-3.13 (m, 9H), 3.02 (dd, J = 11.7 Hz, 4.8 Hz, 1H), 2.82 (m, 1H), 1.38-1.36 (m, 10H), 1.18-1.11 (m, 17H), 1.06 (d, J = 6.8 Hz, 6H).

^{13}C NMR (100 Hz, $CDCl_3$): δ 156.4, 146.7, 143.3, 142.6, 141.5, 140.9, 123.7, 123.6, 123.5, 123.4, 85.1, 79.4, 60.2, 53.7, 43.1, 28.4, 27.9, 27.5, 27.4, 26.9, 24.3, 24.1, 24.0.

HRMS (ESI, pos.) *m/z*: $[M+H]^+$ calcd for $C_{32}H_{52}N_3O_2$, 510.4059; found: 510.4054.

***tert*-butyl ((1,3-bis(2,6-diisopropylphenyl)imidazolidin-4-yl)methyl)carbamate, chloride salt (10)**



10

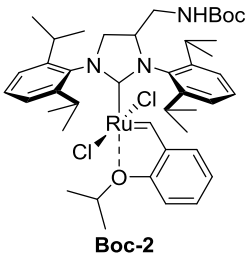
Compound **10** (yield 212 mg, 89 %) was prepared from the same procedure of compound **7**.

^{1}H NMR (400 MHz, $CDCl_3$): δ 8.00 (s, 1H), 7.54-7.46 (m, 3H), 7.32-7.26 (m, 3H), 5.45-5.36 (m, 1H), 5.09 (dd, J = 12.9 Hz, 9.4 Hz, 1H), 4.62 (t, J = 12.1 Hz, 1H), 3.91-3.84 (m, 1H), 3.46-3.41 (m, 1H), 3.23-3.13 (m, 1H), 3.09-2.89 (m, 3H), 1.53 (d, J = 6.8 Hz, 3H), 1.48 (d, J = 6.8 Hz, 3H), 1.40 (dd, J = 6.7 Hz, 0.9 Hz, 6H), 1.36 (s, 9H), 1.26-1.19 (m, 12H).

^{13}C NMR (100 Hz, $CDCl_3$): δ 157.6, 156.7, 146.7, 146.6, 146.3, 145.8, 131.6, 129.1, 127.6, 125.6, 125.2, 125.1, 124.8, 79.2, 65.2, 58.3, 41.4, 29.5, 29.2, 29.1, 29.1, 28.3, 26.0, 25.5, 25.3, 25.2, 23.9, 23.8, 23.7, 23.2.

HRMS (ESI, pos.) *m/z*: $[M-Cl]^+$ calcd for $C_{33}H_{50}N_3O_2$, 520.3903; found: 520.3893.

Boc-2



Boc-2

Boc-2 was prepared from the same procedure of compound **Boc-1**.

A suspension of ^{t}BuOK (67.3 mg, 0.6 mmol) was added to the imidazolinium salt **10** (334 mg, 0.6 mmol) in dry and degassed THF (10 ml). The mixture was stirred 30 min at room temperature. The yellow solution was transferred into a flask containing Hoveyda-Grubbs 1st generation (303 mg, 0.50 mmol) in benzene (40 ml) and stirred at 80 °C. After 1 hour, CuCl (59.4 mg, 0.6 mmol) was slowly added to the solution and stirred for an additional hour at 80°C. The solvent was evaporated the crude

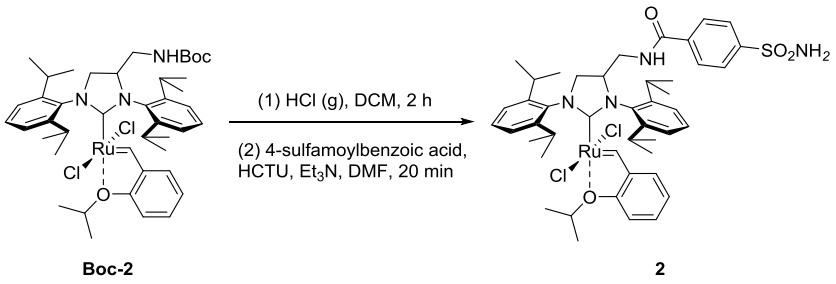
was purified by flash chromatography (10% EtOAc in cyclohexane) yields 107 mg (21%) of product **Boc-2** as a green powder.

^{1}H NMR (400 MHz, CD_2Cl_2): δ 16.12 (s, 1H), 7.57-7.53 (m, 1H), 7.44-7.33 (m, 4H), 7.28-7.23 (m, 2H), 6.80-6.73 (m, 3H), 4.88-4.79 (m, 1H), 4.36-4.19 (m, 3H), 3.53-3.43 (m, 1H), 3.27-3.14 (m, 4H), 1.39-1.10 (m, 39H),

^{13}C NMR (100 Hz, CD_2Cl_2): δ 288.9, 216.6, 164.8, 156.2, 152.7, 149.6, 144.0, 142.6, 137.4, 130.2, 130.1, 129.8, 126.0, 125.6, 124.4, 124.2, 123.8, 122.7, 122.6, 113.4, 79.9, 75.6, 66.6, 58.6, 56.6, 48.2, 43.0, 32.3, 30.1, 29.8, 29.3, 28.6, 28.4, 28.2, 27.2, 26.0, 24.6, 24.1, 23.3, 23.1, 22.1, 21.8, 14.3.

HRMS (ESI, pos.) *m/z*: $[M\text{-}Cl]^+$ calcd for $C_{43}H_{61}ClN_3O_3Ru$, 804.3444; found: 804.3430.

Catalyst **2**



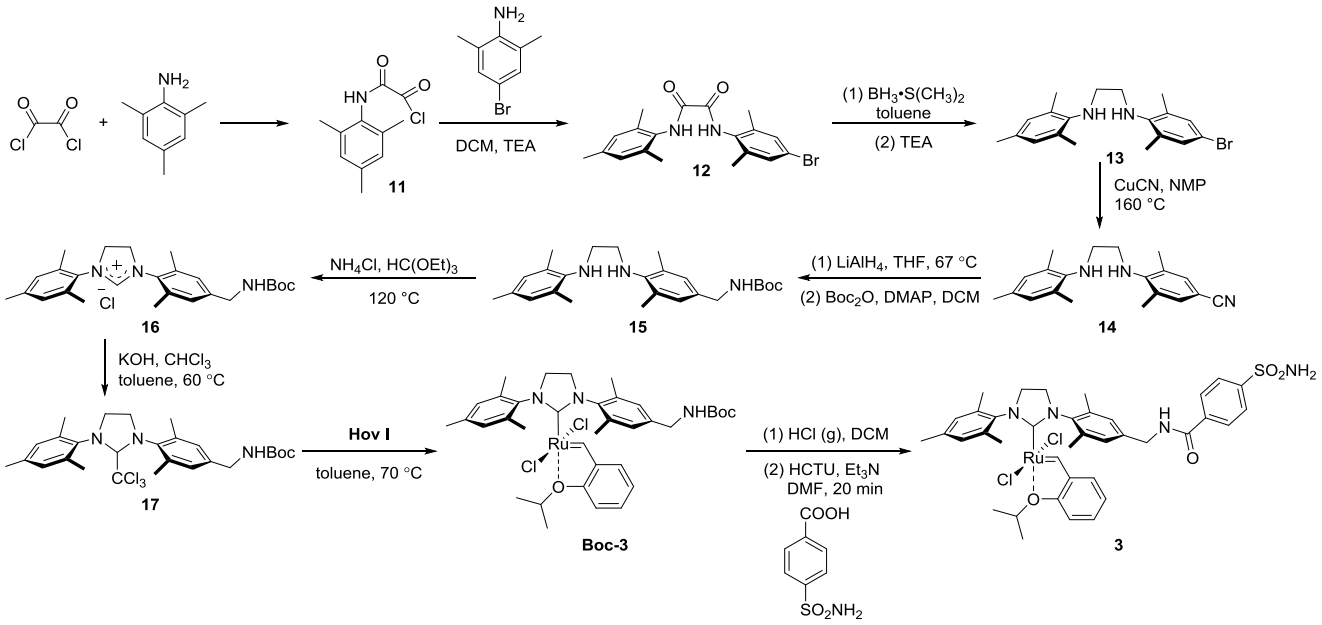
Boc-2 **2**

Catalyst **2** was synthesized according to the preparation procedure used for catalyst **1**. Yield: 13.3 mg, 25%.

^{1}H NMR (400 MHz, CD_2Cl_2): δ 16.1 (s, 1H), 7.90-7.81 (m, 3H), 7.60-7.15 (m, 7H), 6.83-6.57 (m, 4H), 4.94-4.85 (m, 3H), 4.51-4.33 (m, 2H), 3.79-3.34 (m, 4H), 3.22 (br, 1H), 1.35-0.75 (m, 30H).

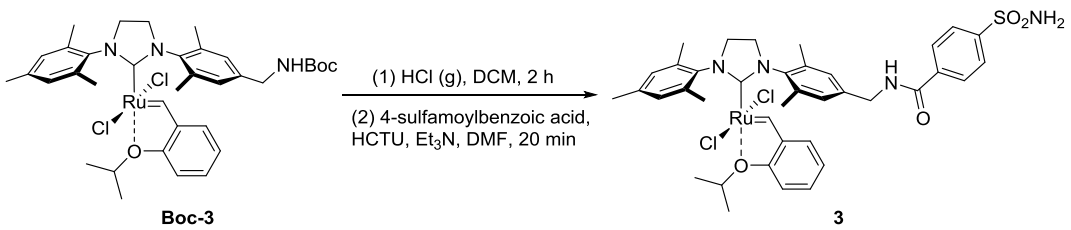
^{13}C NMR (100 Hz, CD_2Cl_2): δ 289.3, 217.5, 166.5, 152.9, 145.7, 144.1, 137.6, 130.6, 130.6, 130.3, 128.7, 127.1, 126.0, 123.0, 122.8, 113.8, 75.9, 41.6, 35.2, 32.5, 32.2, 30.3, 29.6, 27.5, 25.8, 23.2, 22.4, 21.9, 21.0, 19.1, 14.5, 11.8.

HRMS (ESI, pos.) *m/z*: $[M\text{-}Cl]^+$ calcd for $C_{45}H_{58}ClN_4O_4RuS$, 887.2910; found: 887.2915.



Compounds **11**-**17**, **Boc-3** were synthesized according to literature protocols.[5b]

Catalyst **3**



Catalyst **3** was synthesized according to the preparation procedure of catalyst **1**. Yield: 17.4 mg, 21%.

^{1}H NMR (500 MHz, CD_2Cl_2): δ 16.50 (s, 1H), 7.81-7.75 (m, 4H), 7.49-7.46 (m, 2H), 7.24 (s, 2H), 7.06 (s, 2H), 6.91 (d, J = 7.5 Hz, 1H), 6.76 (d, J = 8.3 Hz, 1H), 5.49 (br, 2H), 4.82-4.75 (m, 1H), 4.62-4.61 (m, 2H), 4.12-4.07 (m, 4H), 2.43-2.38 (m, 15H), 1.16 (d, J = 6.2 Hz, 6H).

^{13}C NMR (145 Hz, CD_2Cl_2): δ 296.0, 210.9, 166.2, 152.4, 145.5, 145.4, 138.4, 129.7, 122.6, 129.6, 128.7, 128.4, 128.4, 126.9, 129.9, 122.6, 113.2, 75.7, 52.0, 52.0, 44.0, 21.5, 21.4, 19.8.

HRMS (ESI, pos.) *m/z*: $[M\text{-}Cl]^+$ calcd for $C_{38}H_{44}ClN_4O_4RuS$, 789.1815; found: 789.1807.

Binding affinity for the inhibition of hCA II

A competitive displacement assay was performed according the procedure previously described by Zambel.[8a]

The stock solutions used were phosphate buffer (0.1 M, pH 7.0), dansylamide stock solution (1000 μM, 800 μM, 100 μM, 10 μM) in DMSO, catalyst **2** stock solution (500 μM, 50 μM, 5 μM) in DMSO, and hCA II isoform stock (0.278 μM) in phosphate buffer (0.1 M, pH 7.0).

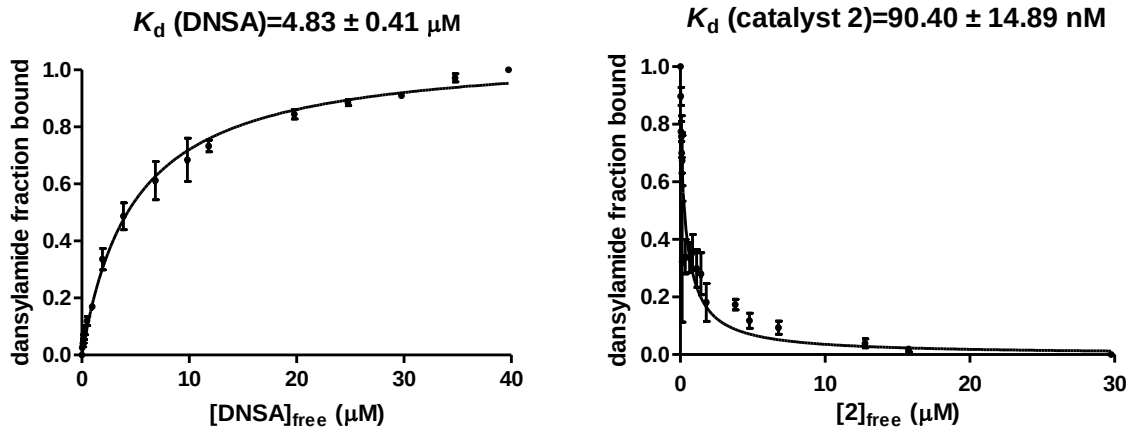
Fluorescence measurements were performed at 29 °C by using TECAN and Nunclon black flat-bottom 96-well plates. The excitation wavelength was set to 280 nm with a fluorescence intensity at 470 nm, and the scan speed was set to 100 nm/min. The K_d value for DNSA was determined by titrating 0.25 μM hCA II (180 μL) with varying concentrations of DNSA (1~20 μL), ranging from 0–40 μM in a total assay volume of 200 μL. The equilibrium dissociation constant for DNSA K_{DNSA} was then determined by fitting the data to equation 1 using Prism 5.0 software.

$$r = \frac{[\mathrm{DNSA}]}{K_{DNSA} + [\mathrm{DNSA}]} \quad \text{(Eq. 1)}$$

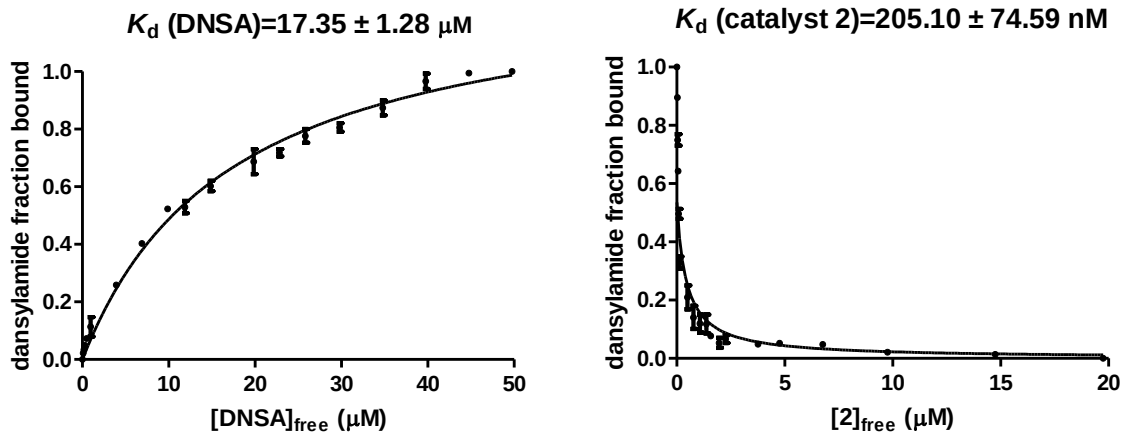
The equilibrium dissociation constants for catalyst **2** were determined by competitive binding with DNSA. A fixed concentration of 20 μM DNSA (5 μL) and 0.25 μM hCA II (180 μL) were then titrated against catalyst **2** from 0–30 μM (1~15 μL). The K_d value of catalyst **2** was then determined by fitting the data to equation 2 with Prism 5.0 software. All the titration experiment were performed in triplicate.

$$r = \frac{1}{1 + (K_{\mathrm{DNSA}}/[\mathrm{DNSA}])(1 + [\mathrm{catalyst}]/K_{\mathrm{d}}(\mathrm{catalyst}))} \quad \text{(Eq. 2)}$$

Determination of DNSA and catalyst **2** dissociation constants (K_d) for hCA II WT.



Determination of DNSA and catalyst **2** K_d for hCA II mutant L198H.



Protein expression and purification

The plasmid encoding human carbonic anhydrase isozyme II (hCA II) and containing a T7 RNA polymerase promoter and an ampicillin resistance gene (pACA) was a generous gift from Carol Fierke, Michigan University. The construct of this plasmid has a serine residue at position 2 instead of an alanine, with no effect on protein expression or catalytic properties.

An overnight pre-culture (50 mL) was used to inoculate 1 L of induction media (20 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, 0.36X M9 salts solution, 0.4%glucose, 60 μM $ZnSO_4$, 100 μg/mL Ampicillin and 34 μg/mL Chloramphenicol). Cells were grown at 37 °C until OD_{600}=0.6-0.8. Protein expression were induced through the addition of IPTG (250 μM final concentration) and $ZnSO_4$ (450 μM final concentration). After 6 to 7 h of incubation at 37 °C at 200 rpm, cells were harvested (4400 rpm, 15 min at 4 °C) and frozen at -20 °C overnight. Cells were lysed by activating the gene encoding T7 lysozyme by three cycles of "freeze/thraw". To the cell pellets, a buffer containing Tris-sulfate (50 mM, pH 8.0), NaCl (50 mM), EDTA (10 mM, pH 8.0), $ZnSO_4$ (0.5 mM), and the protease inhibitors PMSF (10 μg/mL) was added. Cells were resuspended by vigorously shaking at rt for 1 h, followed by adding DNase I (1 μg/L final concentration) for another 1 h shaking until complete digestion of nucleic acids. The cellular remnants were centrifuged (10,000 rpm, 45 min at 4 °C) and cell debris was discarded. After identification of hCA II by SDS-PAGE, the crude supernatant was dialyzed in buffer (50 mM Tris-sulfate pH 8.0 and 0.5 mM $ZnSO_4$) overnight at 4 °C.

The protein was purified by affinity chromatography. Collected fractions were dialyzed in ddH_2O overnight. The protein was lyophilized and kept at 4 °C as white powder.

General procedure for artificial ring closing metathesase of *N*-tosyldiallylamine

A stock solution (200 μM) of the catalyst was prepared by adding DMSO to an aliquot of the catalyst. A stock solution of *N*-tosyldiallylamine (20 mM) was prepared in a separate vial. A stock solution of protein (13.3 μM) was prepared by using 0.1 M phosphate buffer pH 7.0.

In a small glass reaction vial, protein stock solution (180 μL) and catalyst DMSO stock solution (10 μL) were added and incubated at 37 °C for 20 min. Then substrate *N*-tosyldiallylamine stock solution (10 μL) was added, and the reaction vial was placed in an incubator for 4 h at 37 °C.

Upon completion of the reaction, the mixture was added internal standard 2-phenylethanol (100 μL, 1 mM water solution) and MeOH (700 μL). The solution mixture was transferred to an eppendorf tube and centrifuged at 14'000 rpm for 15 minutes to precipitate the protein. The supernatant (500 μL) was transferred in an HPLC vial with additional 500 μL water and the sample was subjected to RP-HPLC to determine the TON and conversion.

HPLC analysis. Column: XDB-C18, Eclipse by Agilent: 150 x 4.6 mm; 5 μm with guard column. Method: $V_{injected}$: 50 μL. Eluent (Solvent A: H_2O) , (Solvent B: CH_3CN) 10% B at 0 min, 10% B at 5 min; 90% B at 20 min; 10 % B at 30 min. Detection at 210 nm. T_R 12.2 min (internal standard 2-phenylethanol), 14.6 min (product 1-tosyl-2,5-dihydro-1*H*-pyrrole), 16.1 min (substrate *N*-tosyldiallylamine).

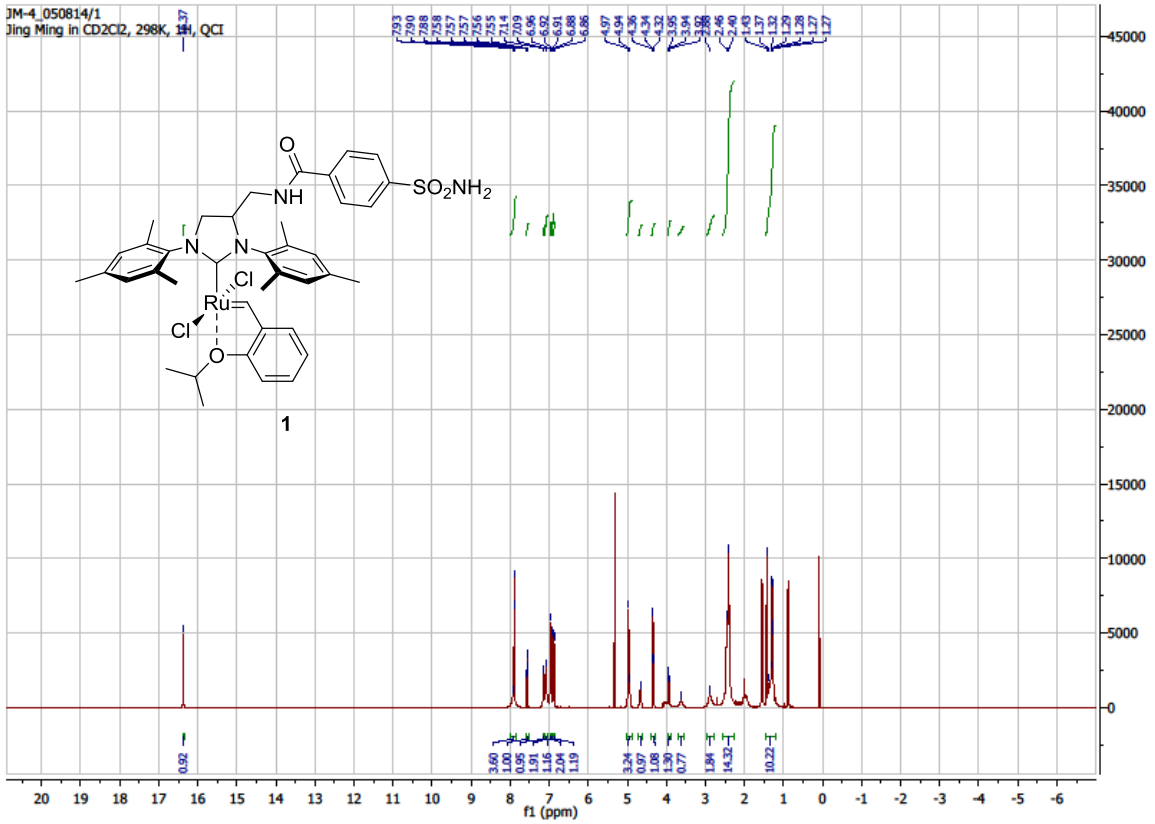
Docking simulation results

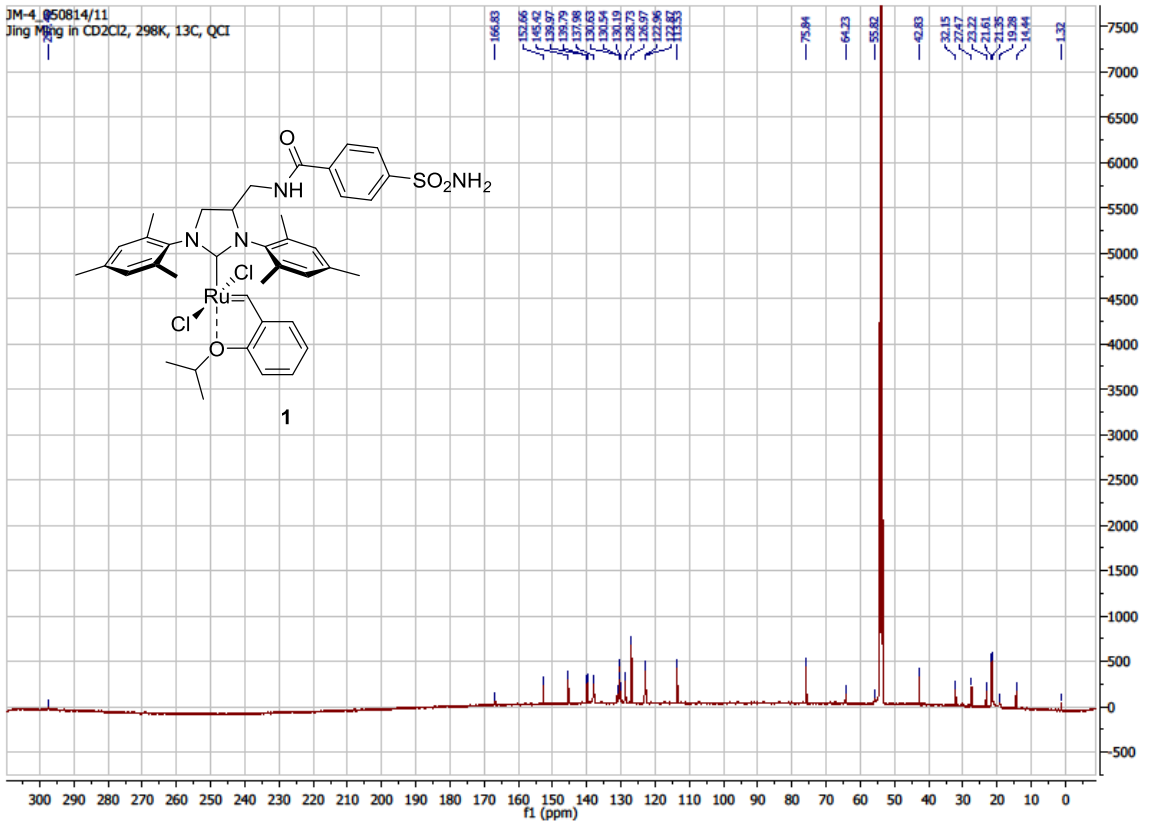
Both (*R*) and (*S*)-enantiomers of catalyst **2** were docked within hCA II WT by GOLD programme. The following table summarizes the calculation results of the best fitted catalyst within hCA II.

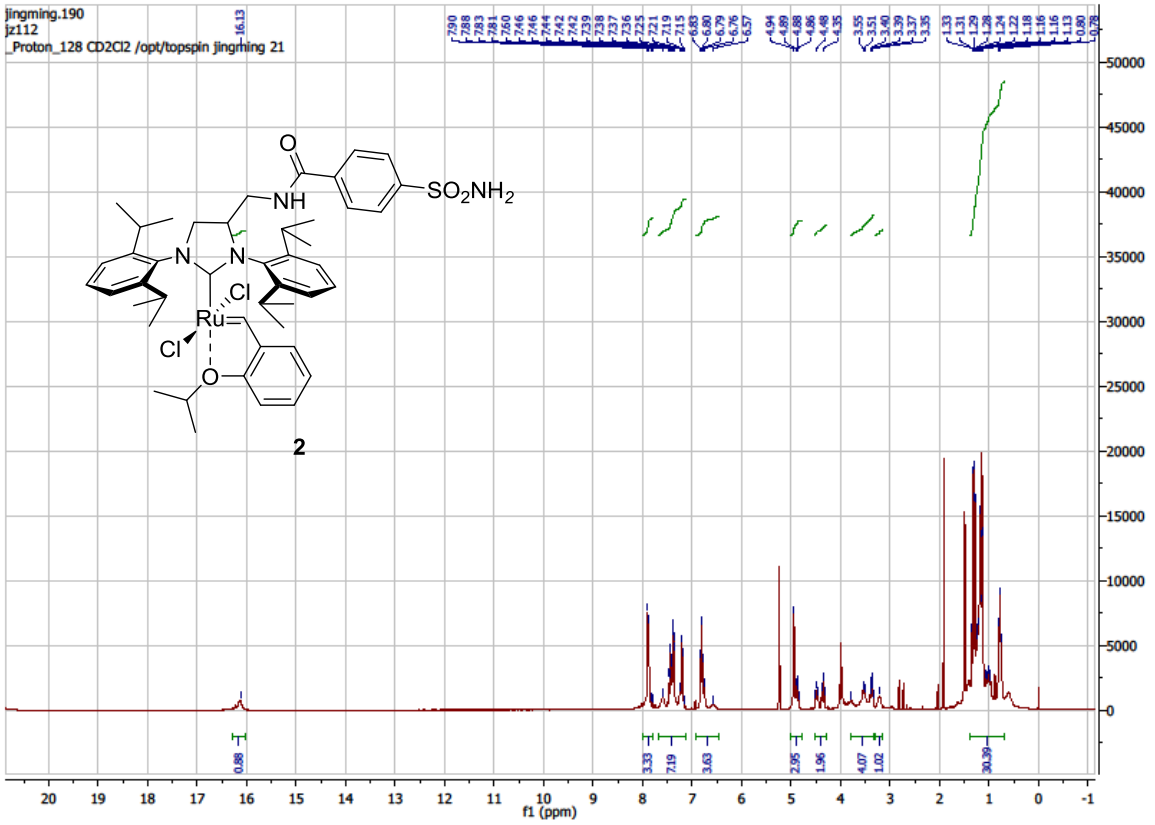
Host protein	Score*	ΔG^+	S_{hbond}*	S_{lipo}*	H_{rot}*	ΔE_{clash}*	ΔE_{int}*
(*R*)-**2** ⊂ WT hCA II	32.77	-26.92	1.25	187.00	1.80	-8.55	0.84
(*S*)-**2** ⊂ WT hCA II	40.86	-33.03	1.29	238.03	1.80	-11.51	1.91

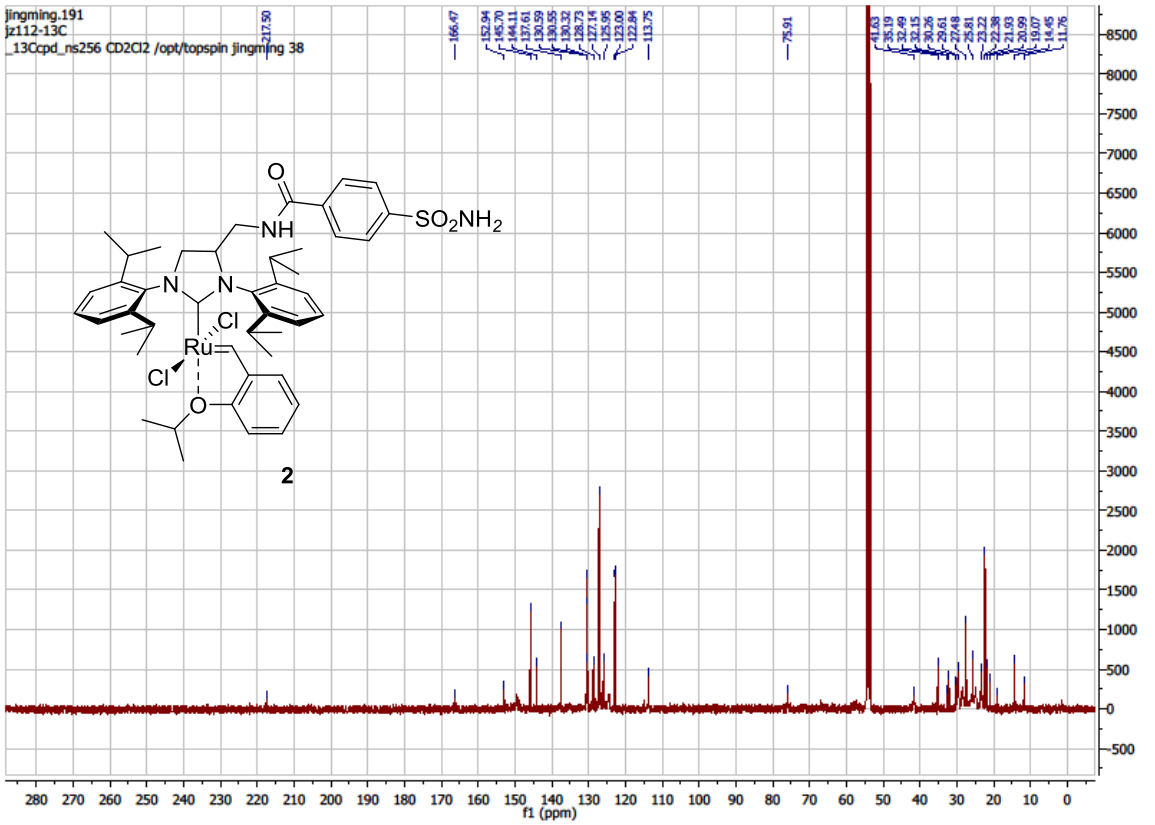
* Values dimensionless $^+$ Values in $kJ \cdot mol^{-1}$

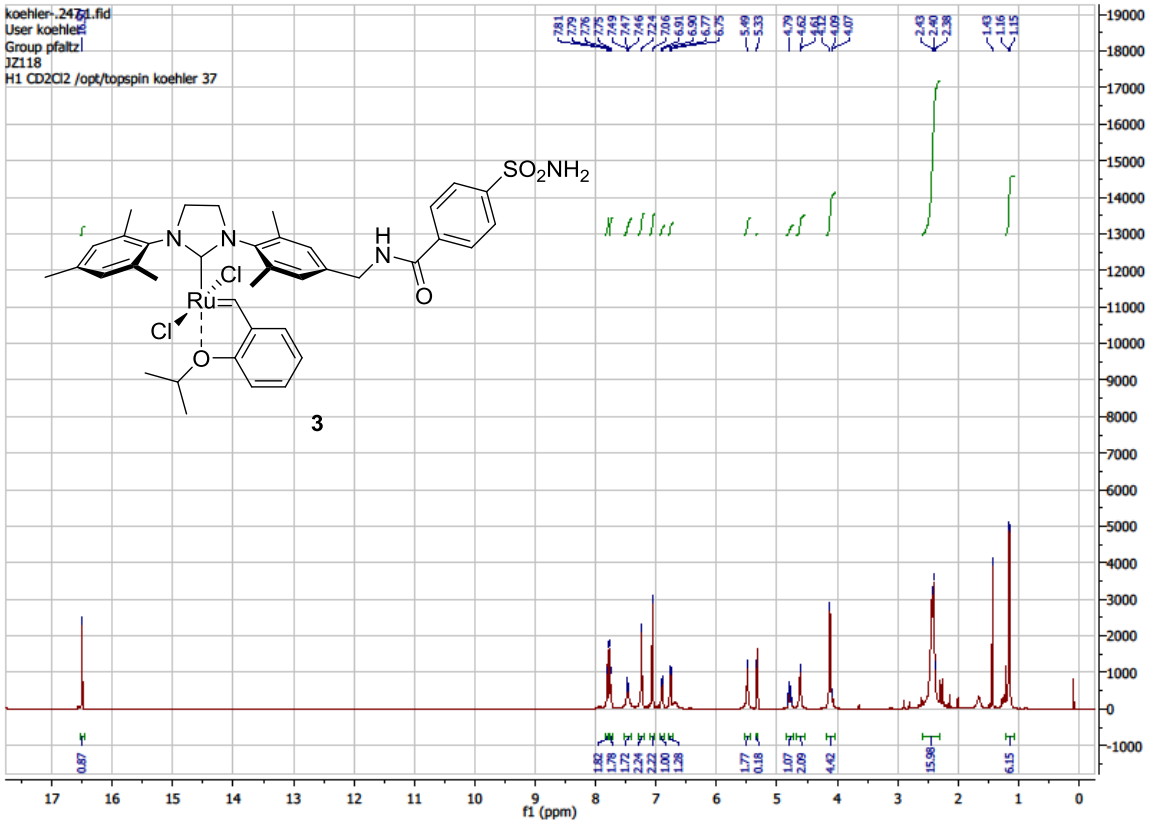
NMR spectra:

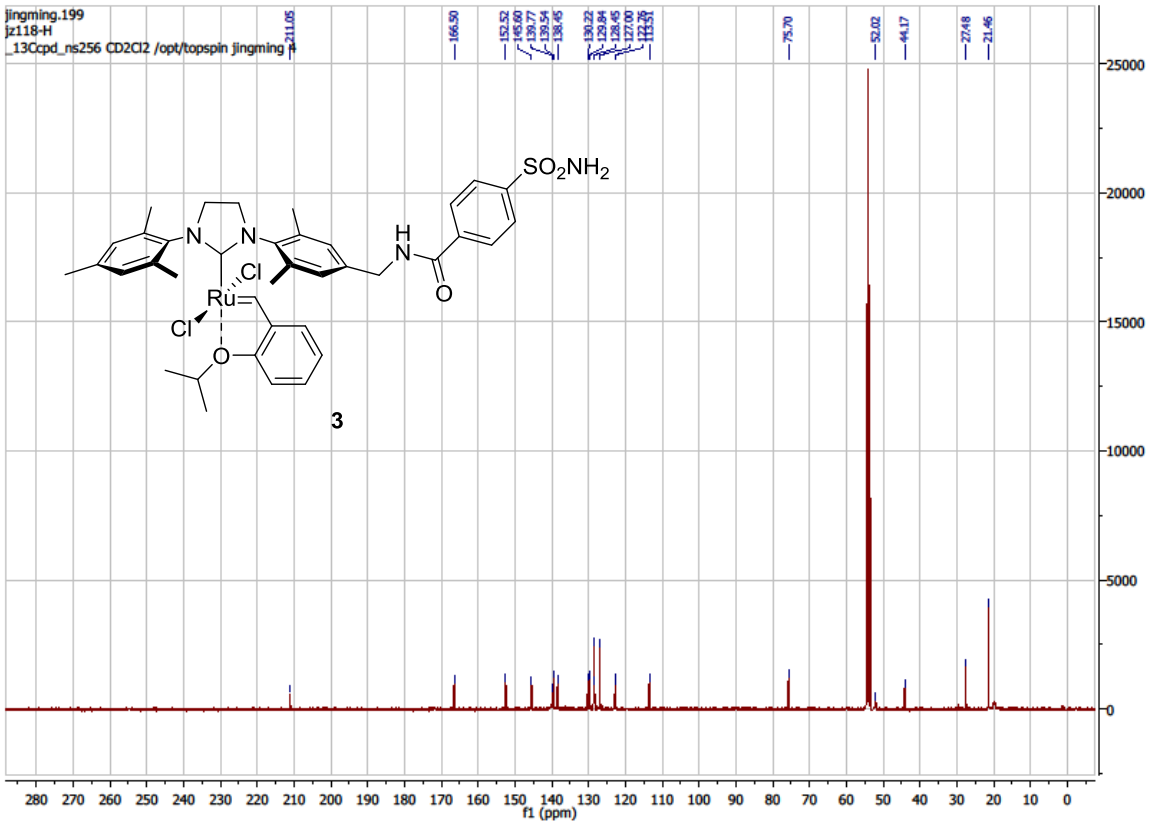

JM-4_050814/1
Jing Ming in CD2Cl2, 298K, 1H, QCI
SO2NH2
NH
Cl
Ru
O
1
f1 (ppm)


JM-4_050814/11
Jing Ming in CD2Cl2, 298K, 13C, QCI
SO2NH2
NH
Cl
Ru
O
1
f1 (ppm)


jingming.190
jz112
_Proton_128 CD2Cl2 /opt/topspin jingming 21
O
NH
SO2NH2
N
N
Cl
Ru
Cl
O
2
f1 (ppm)


jingming.191
jz112-13C
_13Ccpd_ns256 CD2Cl2 /opt/topspin jingming 38
O
NH
SO2NH2
N
N
Cl
Ru
Cl
O
2
f1 (ppm)


koehler-247-1.fid
User koehler
Group pfaltz
JZ118
H1 CD2Cl2 /opt/topspin koehler 37
SO2NH2
N
N
H
N
Cl
Ru
Cl
O
O
3
f1 (ppm)


jingming.199
jz118-H
_13Ccpd_ns256 CD2Cl2 /opt/topspin jingming
SO2NH2
N
N
H
N
Cl
Ru
Cl
O
O
3
f1 (ppm)

2.6 Some comments on human carbonic anhydrase II

Some comments and suggestions are coming as follows:

(a) Previous research has highlighted that the histidine imidazoles of the active site are beneficial biological ligands by coordination to various transition metals to provide unforeseen biochemical activity. Accordingly, metal substitution remains a promising strategy to create new ArMs based on hCA II. It is just a question of whether the newly formed metalloprotein is stable, whether the metal-protein binding affinity is high and whether the second coordination sphere may effectively influence the metal's activity. Also, the coordination of metal ions to the abundant hCA II surface histidine residues should be minimized to minimize side reactions.

(b) New catalytic transition metal catalysts bearing arylsulfonamide anchors need to be synthesized for the exploration of new reaction. Here, the catalysts should be active and stable enough under physioligical conditions. As multistep synthesis and purification of metal complex (especially the ligand screening for the 1st coordination site optimization) are time-consuming, it is advisable to build cooperation with organic and organometallic research groups for new reaction development.

(c) Bioconjugation or ligation with new metal ions to hCA II scaffold is another potential method to generate new ArMs. For example, two popular methods are Cu^{+} catalyzed azide-alkyne cycloaddition and thiol Michael addition under biorthogonal conditions to cross link biomolecules.[10]

(d) There are only two reported examples for directed evolution of carbonic anhydrase: one is about esterase activity improvement[11] and the other concerns carbon capture from flue gas[12]. The directed evolution of carbonic anhydrase for new transition metal catalysis is still unexplored. Evolving carbonic anhydrase in a high throughput manner could be a promising topic in future years.

2.7 Reference

(1) (*a*) F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum and V. L. Pecoraro, *Chem. Rev.*, 2014, **114**, 3495; (*b*) J. C. Lewis, *ACS Catal.*, 2013, **3**, 2954; (*c*) A. Ilie and M. T. Reetz, *Isr. J. Chem.*, 2014, **55**, 51; (*d*) M. Dürrenberger and T. R. Ward, *Curr. Opin. Chem. Biol.*, 2014, **19**, 99; (*e*) A. J. Boersma, R. P. Megens, B. L. Feringa and G. Roelfes, *Chem. Soc. Rev.*, 2010, **39**, 2083; (*f*) Y. Lu, N. Yeung, N. Sieracki and N. M. Marshall, *Nature*, 2009, **460**, 855; (*g*) P. J. Deuss, R. den Heeten, W. Laan and P. C. Kamer, *Chem. - Eur. J.*, 2011, **17**, 4680.

(2) N. Krall, F. Pretto,W. Decurtins, G. J. L. Bernardes, C. T. Supuran and D. Neri, *Angew. Chem., Int. Ed.*, 2014, **53**, 4231.

(3) (*a*) C. Lo, M. R. Ringenberg, D. Gnandt, Y. Wilson and T. R. Ward, *Chem. Commun.*, 2011, **47**, 12065; (*b*) C. Mayer, D. G. Gillingham, T. R. Ward and D. Hilvert, *Chem. Commun.*, 2011, **47**, 12068; (*c*) T. Matsuo, C. Imai, T. Yoshida, T. Saito, T. Hayashi and Shun Hirota, *Chem. Commun.*, 2012, **48**, 1662; (d) D. F. Sauer, M. Bocola, C. Broglia, M. Arlt, L. Zhu, M. Brocker, U. Schwaneberg and J. Okuda, *Chem. - Asian J.*, 2015, **10**, 177; (*e*) F. Philippart, M. Arlt, S. Gotzen, S. J. Tenne, M. Bocola, H. H. Chen, L. Zhu, U. Schwaneberg and J. Okuda, *Chem. - Eur. J.*, 2013, **19**, 13865; (*f*) D. Burtscher and K. Grela, *Angew. Chem., Int. Ed.* 2009, **48**, 442; (f) Y. Lin, J. M. Chalker and B. G. Davis, *ChemBioChem*, 2009, **10**, 959.

(4) (*a*) V. M. Krishnamurthy, G. K. Kaufman, A. R. Urbach, I. Gitlin, K. L. Gudiksen, D. B. Weibel and G. M. Whitesides, *Chem. Rev.*, 2008, **108**, 946; (*b*) F. W. Monnard, E. S. Nogueira, T. Heinisch, T. Schirmer and T. R. Ward, *Chem. Sci.*, 2013, **4**, 3269; (*c*) M. Schmid, E. S. Nogueira, F. W. Monnard, T. R. Ward and M. Meuwly, *Chem. Sci.*, 2012, **3**, 690; (*d*) F. W. Monnard, T. Heinisch, E. S. Nogueira, T. Schirmer and T. R. Ward, *Chem. Commun.*, 2011, **47**, 8238; (*e*) D. Can, B. Spingler, P. Schmutz, F. Mendes, P. Raposinho, C. Fernandes, F. Carta, A. Innocenti, I. Santos, C. T. Supuran and R. Alberto, *Angew. Chem., Int. Ed.*, 2012, **51**, 3354.

(5) (*a*) J. P. Jordan and R. H. Grubbs, *Angew. Chem., Int. Ed.*, 2007, **46**, 5152; (*b*) A. Kajetanowicz, A, Chatterjee, R, Reuter and T. R. Ward, *Catal. Lett.*, 2014, **144**, 373.

(6) T. R. Ward, *Acc. Chem. Res.*, 2011, **44**, 47.

(7) V. M. Robles, M. Dürrenberger, T. Heinisch, A. Lledós, T. Schirmer, T. R. Ward and J. D. Maréchal, *J. Am. Chem. Soc.*, 2014, **136**, 15676.

(8) (*a*) S. C. Wang and D. B. Zamble, *Biochem. Mol. Biol. Educ.*, 2006, **34**, 364. (*b*) A. L. Banerjee, M. Swanson, B. C. Roy, X. Jia, M. K. Haldar, S. Mallik and D. K. Srivastava, *J. Am. Chem. Soc.*, 2004, **126**, 10875.

(9) (a) M. D'hooghe, I. Kerkaert, M. Rottiers and N. D. Kimpe, *Tetrahedron*, 2004, **60**, 3637; (b) J. P. Jordan and R. H. Grubbs, *Angew. Chem., Int. Ed.*, 2007, **46**, 5152.

(10) J. Dadová, M. Vrábel, M. Adámik, M. Brázdová, R. Pohl and M. Fojta, *Chem. - Eur. J.*, 2015, **21**, 16091.

(11) S. M. Gould and D. S. Tawfik, *Biochemistry*, 2005, **44**, 5444.

(12) O. Alvizo, L. J. Nguyen, C. K. Savile, J. A. Bresson, S. L. Lakhapatri, E. O. P. Solis, R. J. Fox, J. M. Broering, M. R. Benoit, S. A. Zimmerman, S. J. Novick, J. Liang and J. J. Lalonde, *Proc. Natl. Acad. Sci.* 2014, **111**, 16436.

Chapter 3 | An artificial metalloenzyme for carbenoid transfer based on a biotinylated dirhodium anchored within streptavidin

Jingming Zhao,[a] Daniel G. Bachmann,[b] Markus Lenz,[c] Dennis G. Gillingham[b] and Thomas R. Ward*[a]

[a]Department of Chemistry, University of Basel, Mattenstrasse 24a, BPR 1096, Basel, CH-4058, Switzerland.

[b]Department of Chemistry, University of Basel, St. Johanns-Ring 19, CH-4056, Basel, Switzerland.

[c]Institute for Ecopreneurship, School of Life Sciences, University of Applied Sciences and Arts Northwestern Switzerland, Gründenstrasse 40, CH-4132 Muttenz, Switzerland.

This work was submitted and under review.

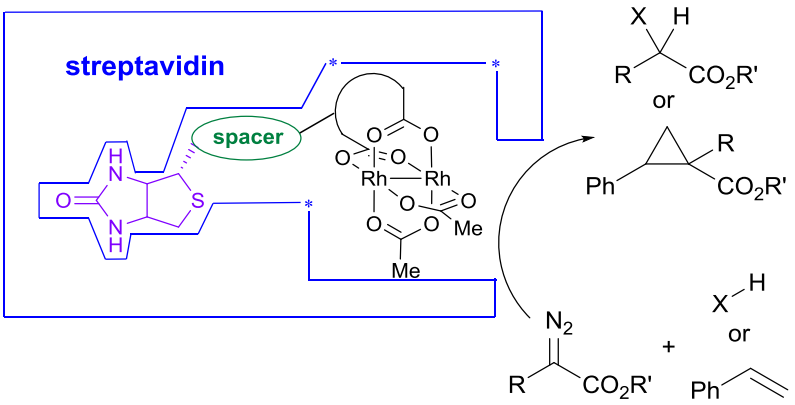
3.1 Abstract

We report on artificial metalloenzymes that incorporate a biotinylated dirhodium core embedded within engineered streptavidin variants. The resulting biohybrid catalyzes the carbenoid insertion in C–H bonds and olefins. Chemical- and genetic optimisation allows to modulate the catalytic activity of the artificial metalloenzymes that are shown to be active in the perisplasm of *E. coli* (up to 20 turnovers).

3.2 Introduction

Dirhodium(II) tetracarboxylate complexes have been shown to be exceptionally active catalysts for carbene transfer reactions including cyclopropanation[1-3] and X-H insertion (X= C, N, O, S, Si etc)[4-8]. Thanks to their robustness, remarkable activity and selectivity under physiological conditions, they have also found applications as a versatile tool in chemical biology. For this purpose, dirhodium-tetracarboxylate moieties have been linked to biomacromolecules including peptides, proteins and oligonucleotides. The Ball group developed dirhodium metallopeptides for protein modification and intracellular imaging.[9,10] Gillingham and coworkers have reported dirhodium complexes for DNA modification[11], aqueous catalysis and metal uptake within tumor cells[12]. In the context of artificial metalloenzymes (ArMs), the Lewis group covalently linked a dirhodium tetracarboxylate bearing a terminal alkyne to a genetically-engineered prolyl oligopeptidase equipped with an azidophenylalanine. The resulting ArM displayed excellent catalytic properties for intermolecular cyclopropanation. Relying on directed evolution (either random or targeted), both the activity and the enantioselectivity could be significantly improved.[13,14] The evolved dirhodium cyclopropanase was further shown to catalyze other carbene transfer reactions.

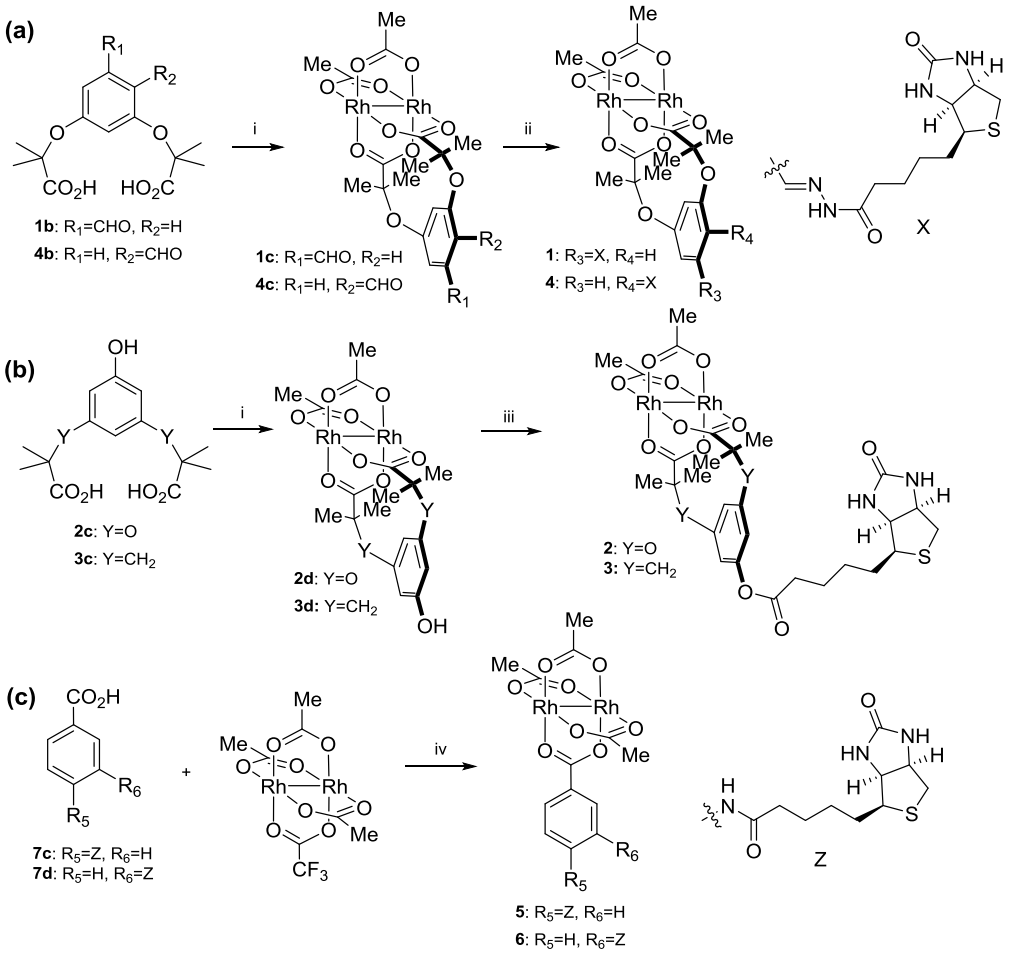
Pioneered by Wilson and Whitesides in 1978, artificial metalloenzymes based on the biotin-(strept)avidin technology, have been implemented for a large variety of reactions including: (transfer)-hydrogenation[15,16], cross-coupling[17], metathesis etc.[18–20] Since then, many groups have reported on the creation of various ArMs relying on alternative anchoring strategies with promising catalytic properties.[21–28] Building upon our experience with ArMs based on the biotin-streptavidin technology relying on precious metal cofactors,[29,30] we present herein our efforts to anchor a bulky biotinylated dirhodium moiety within engineered streptavidin to catalyze carbenoid transfer reactions, Scheme 1.



Scheme 1. Artificial metalloenzymes for carbenoid transfer based on a dirhodium tetracarboxylate moiety. The catalytic properties of the ArM can be chemo-genetically optimized: variation of the spacer (green) between the biotin anchor and the dicarboxylate moiety can be combined with the introduction of point mutations on the streptavidin scaffold (blue stars).

3.3 Results and discussion

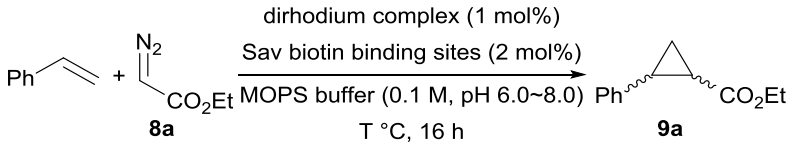
Initially, we designed four bis-chelating dirhodium cofactors **1**-**4** (Scheme 2). Substitution of either two acetate- or trifluoracetate ligands from $Rh_2(OAc)_4$ and (*cis*)-$Rh_2(OAc)_2(OCOCF_3)_2$[31] respectively with the *m*-substituted dicarboxylate-bearing ligands **1b**, **2c**, **3c**, **4b** yielded the dirhodium intermediates **1c**, **2d**, **3d**, **4c** respectively (see ESI for experimental details). Biotinylation using biotin derivatives **7a** or **7b**, through either a hydrazine coupling or an esterification afforded the target biotinylated dirhodium cofactors **1**-**4**. In light of the remarkable inertness of the $Rh_2(\mu^2{-}O_2CR)_4$ moiety, we also prepared the dirhodium complexes **5** and **6**. For this purpose, $Rh_2(OAc)_3(OCOCF_3)$[31] was treated with biotinylated monocarboxylic acid **7c** or **7d** to afford the corresponding complexes **5** and **6** respectively. All biotinylated dirhodium complexes were purified by reversed-phase preparative HPLC and fully characterised (See SI).



Scheme 2. Biotinylated dirhodium tetracarboxylate complexes tested in this study: (a) catalyst **1** and **4**; (b) catalyst **2** and **3**; (c) catalyst **5** and **6**. Reaction conditions: i) $Rh_2(OAc)_4$, *N*, *N*-dimethylaniline, 140 °C, 3 h or (*cis*)-$Rh_2(OAc)_2(OCOCF_3)_2$, K_2CO_3, THF, 50 °C, 3 h; ii) (+)-biotin hydrazide **7a**, TFA, DMSO, rt, 3 h; iii) biotin pentafluoro phenyl ester **7b**, NaH, DMSO, rt, 3 h; iv) *N*, *N*-Diisopropylethylamine, DMSO, 50 °C, 2 h.

In the absence of Sav, the dirhodium complexes **1**-**6** (25 ^{o}C, pH 7.0, MOPS buffer (0.1 M) outperform (i.e. higher turnover number, TON) $Rh_2(OAc)_4$ for the cyclopropanation of styrene Table 1, entries 1-7). Incorporation of the dirhodium cofactors **1**-**6** into wild-type streptavidin (Sav WT) leads to an erosion in activity (Table 1, entries 9-14). Cofactor **2** outperforms all other cofactors: **2** · Sav WT affords 54 TON, compared to 79 TON in the absence of Sav. The biotinylated monodentate dirhodium complexes **5** and **6** lead to significantly lower TONs when embedded with Sav. We hypothesize that this may be due to the loss of the biotinylated ligand accompanied by decomposition of the dirhodium moiety. Increasing or reducing the temperature leads to an erosion in TON for **2** · Sav WT (Table 1, entries 15, 16). Screening at various pH highlights that the dirhodium cofactor **2** performs best at neutral pH (Table 1, entries 17-20). Based on this initial screen, cofactor **2** was selected for all further investigations.

Table 1 Selected results for the ArMs-catalysed cyclopropanation of styrene with ethyl diazoacetate **8a**.



Entry	Catalyst	Sav	Tem (°C), pH	TON[a]	trans/cis
1	$Rh_2(OAc)_4$	–	25, 7.0	2 ± 0	1.5/1
2	**1**	–	25, 7.0	68 ± 1	1.4/1
3	**2**	–	25, 7.0	79 ± 0	1.3/1
4	**3**	–	25, 7.0	79 ± 1	1.4/1
5	**4**	–	25, 7.0	52 ± 1	1.3/1
6	**5**	–	25, 7.0	60 ± 1	1.4/1
7	**6**	–	25, 7.0	46 ± 1	1.4/1
8	$Rh_2(OAc)_4$	WT	25, 7.0	2 ± 0	1.5/1
9	**1**	WT	25, 7.0	34 ± 3	1.3/1
10	**2**	WT	25, 7.0	54 ± 7	1.4/1
11	**3**	WT	25, 7.0	38 ± 2	1.4/1
12	**4**	WT	25, 7.0	33 ± 1	1.3/1
13	**5**	WT	25, 7.0	3 ± 0	1.1/1
14	**6**	WT	25, 7.0	1 ± 0	1.3/1
15	**2**	WT	5, 7.0	4 ± 1	1.2/1
16	**2**	WT	55, 7.0	0	–
17	**2**	WT	25, 4.0[c]	2 ± 0	1.3/1
18	**2**	WT	25, 5.0[d]	29 ± 4	1.4/1
19	**2**	WT	25, 6.0	50 ± 6	1.4/1
20	**2**	WT	25, 8.0	12 ± 4	1.4/1

[a] Reaction conditions: [styrene] = 10 mM, [**8a**] = 30 mM, [catalyst] = 100 μM, [Sav biotin binding sites] = 200 μM, V_{tot} = 400 μL (5% DMSO), 25 °C for 16 hours. [b] TON = turnover number determined by GC using 1, 3, 5-trimethoxybenzene as internal standard. [c] acetate buffer (0.1 M, pH 4.0). [d] acetate buffer (0.1 M, pH 5.0).

For genetic optimization purposes, we screened purified Sav mutants for both cyclopropanation and C-H insertion reactions in the presence of cofactor **2**. The results are summarized in Figure 1a using ethyl diazoacetate **8a** and Figure 1b using donor-acceptor ethyl diazo(phenyl)acetate **8b** respectively.

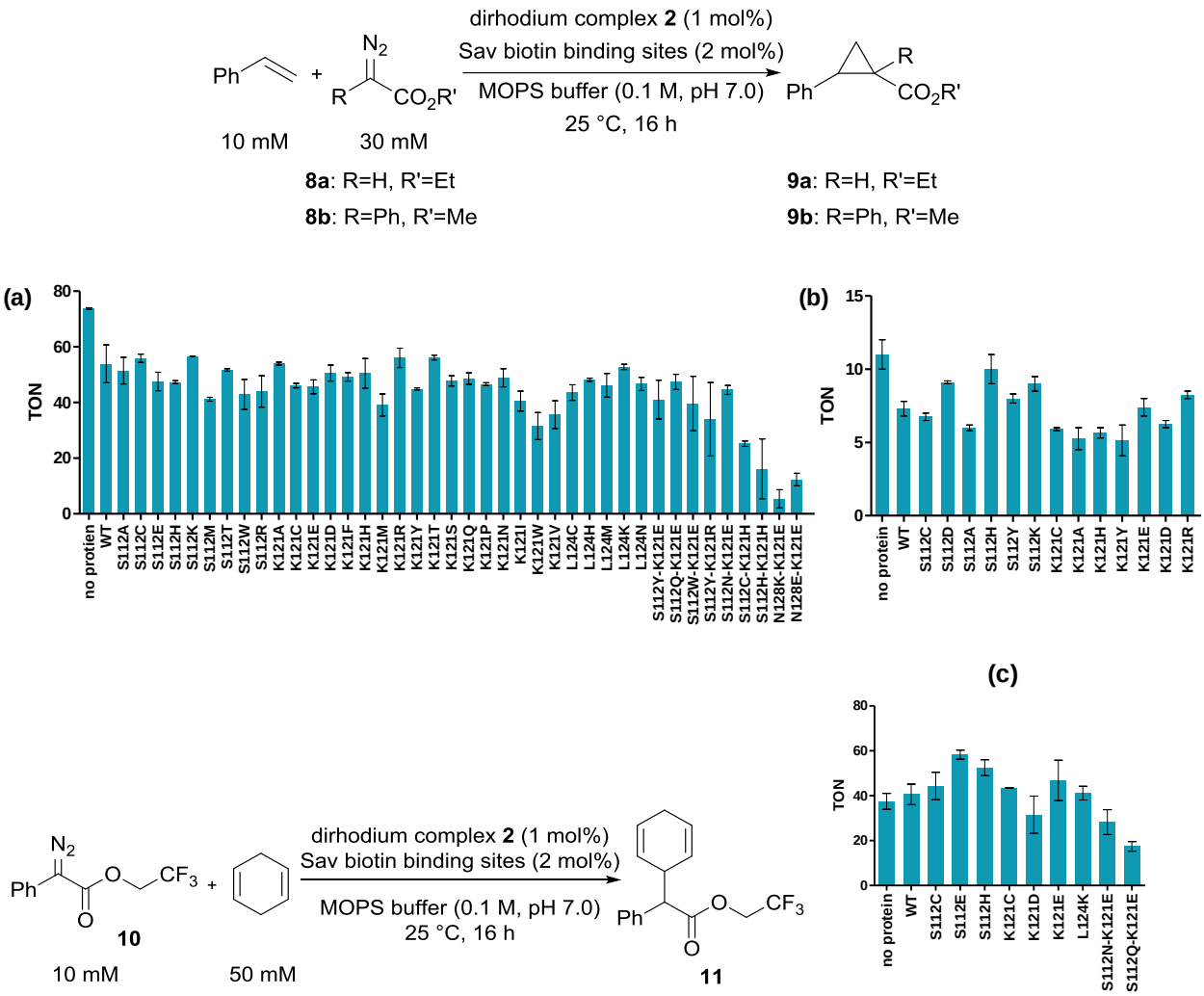


Figure 1. Experimental conditions and selected results for the ArM's catalysed cyclopropanation of styrene with ethyldiazoacetate **8a** (a) and diazo(phenyl)acetate **8b** (b), and C-H insertion of diazo **10** with 1,4-cyclohexadiene (c).

Next, we investigated the C-H insertion of trifluoroethyl (phenyl)diazoacetate **10** with 1, 4-cyclohexadiene catalysed by **2** · Sav and mutants thereof. The reaction afforded exclusively the allylic insertion product **11**, with no double insertion or cyclopropanation byproducts detected. A selection of Sav mutants were screened. Gratifyingly, **2** · Sav S112E and K121E outperformed both the free cofactor **2** and **2** · Sav WT. No enantioselectivity could be detected however.

To gain structural insight into the localisation of biotinylated dirhodium complex within Sav WT, a docking simulation for the hydrated dirhodium complex **2**·$(H_2O)_2$ within Sav WT was performed using the GOLD software suite, Figure 2.[32] The minimized docked structure suggests that the dirhodium moiety protrudes out of the biotin-binding vestibule: The shortest C_α-Rh distances are: 6.1 Å, 6.0 Å, 7.4 Å for the closest lying amino acids S112, K121 and

L124 respectively, Figure 2. This may explain why the genetic optimization has such a modest effect on the performance of the ArM.

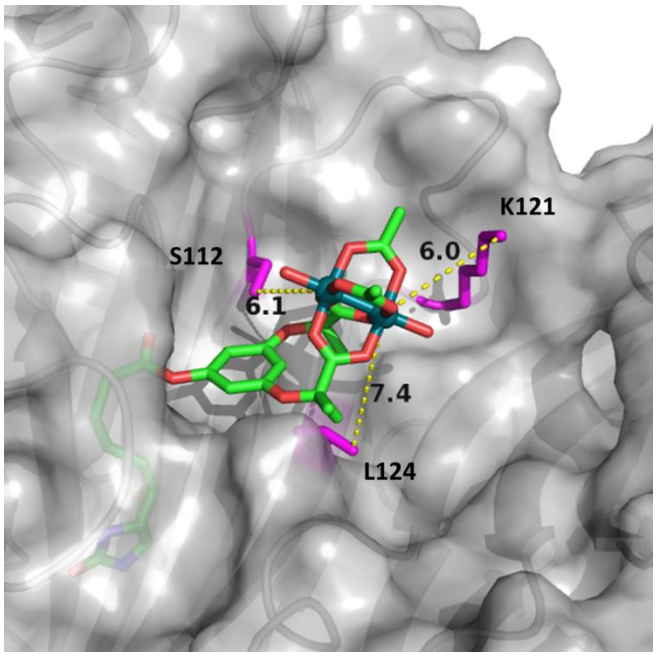


Figure 2. Structure resulting from a docking simulation of dirhodium complex **2**·$(H_2O)_2$ · Sav WT. The protein is displayed as light grey surface and the biotinylated cofactor **2** is displayed as color-coded sticks. The closest lying amino acids are displayed as magenta sticks. The closest contacts between the cofactor and S112, K121 and L124 are highlighted with yellow dotted lines.

It has been reported that dirhodium tetracarboxylate moieties tolerate cellular components.[33] We thus secreted the Sav to the periplasm and screened the artificial cyclopropanase in the presence of whole *E. coli* cells. Figure 3 outlines the protocol implemented for the periplasmic screening.[34] The TOP10(DE3)_pET30 strain was used to express and secrete Sav into the periplasm. The following isoforms were tested: Sav WT, S112C, S112D, K121C and K121M. After harvesting, the cell pellets were incubated with a MOPS buffer containing 50 µM dirhodium cofactor **2** for 30 min on ice. The unbound cofactor **2** was washed away and the catalysis buffer containing the substrate was added to the cell pellet. Gratifyingly, the artificial cyclopropanase **2** · Sav WT outperformed the free cofactor **2**. Introduction of a cysteine residue, either at position S112C or K121C leads to improved activities. ICP-MS analysis revealed significant Rh-accumulation within *E. coli* harboring periplasmic Sav (Sav^{peri}): *E. coli* not secreting Sav (empty *E. coli*) contained 3.2 nmol Rh (corresponding to 6 % uptake from the 50 nmol Rh-added to the supernatant). For *E. coli* harboring Sav^{peri}, the Rh-amounts determined varied between 5.6 − 8.9 nmol (i.e. 11% − 18% uptake) depending on the mutant. Accordingly, the TONs are 10 for the empty *E. coli*, 9 for Sav^{peri} WT, 17 for Sav^{peri} S112C and 20 for Sav^{peri} K121C respectively (See SI).

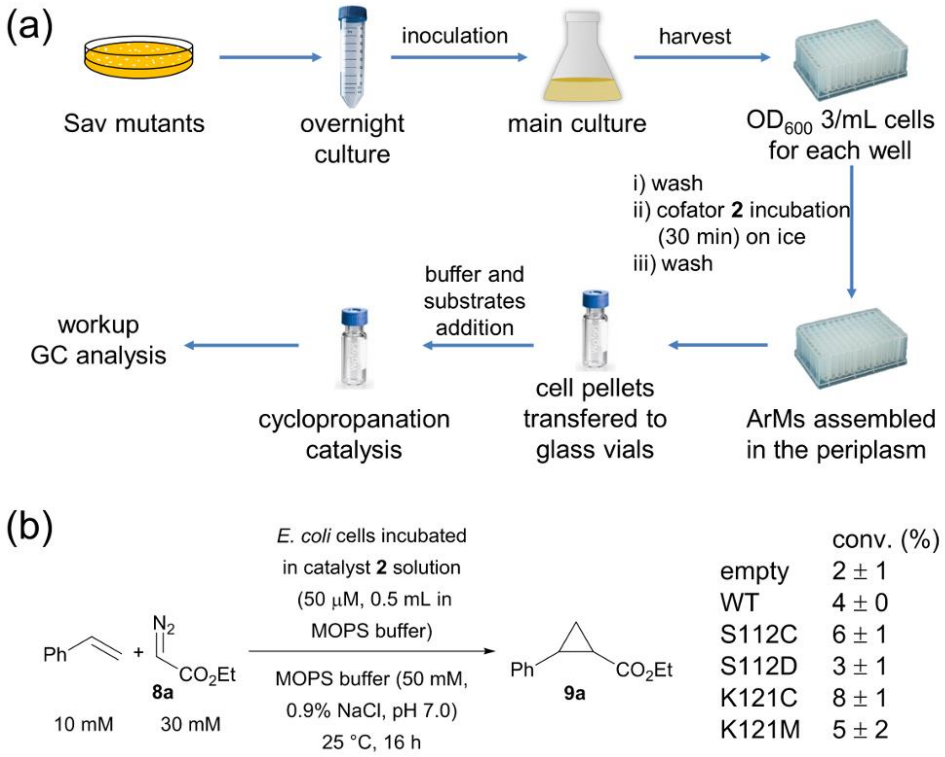


Figure 3. An *in cellulo* artificial cyclopropanase workflow (a), reaction conditions and screening results (b) for the periplasmic screening.

3.4 Conclusion

In summary, we have developed an artificial carbenoid transferase based on the biotin-streptavidin technology. The resulting ArMs catalyse both intermolecular cyclopropanation and C-H insertion. The dirhodium ArMs showed their ability to maintain their activity in the periplasm of *E. coli* cells, highlighting the robustness of the dirhodium moiety in a cellular environment. This paves the way for high-throughput screening to further optimize the activity and the selectivity of such artificial carbenoid-transferases based on the biotin-streptavidin technology. In this context, we have recently reported on chimeric Sav that incorporate extended loops around the biotin-binding vestibule.[35] We hypothesize that these chimeras may offer a better-defined environment around the bulky dirhodium cofactor **2**, thus offering a means to influence the selectivity of the reactions.

3.5 Supporting information

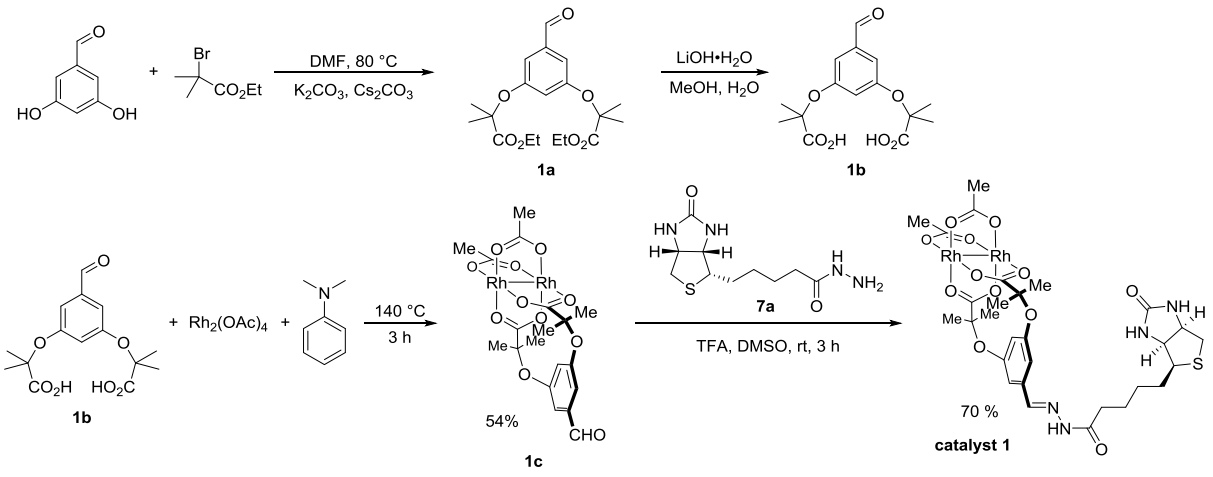
General aspects

Materials and reagents were purchased from the highest commercially available grade and used without further purification.

^{1}H and ^{13}C spectra were recorded on a Bruker 400 MHz, 500 MHz and 600 MHz. Chemical shifts are reported in ppm (parts per million). Signals are quoted as s (singlet), d (doublet), t (triplet), brs (broad) and m (multiplet). High resolution ^{1}H, ^{13}C and 2D spectra (for final biotinylated dirhodium complexes) were recorded on a 600 MHz Avance III equipped with a QCI probe-head. Electron-Spray Ionization Mass Spectra (ESI-MS) were recorded on a Bruker FTMS 4.7T bioAPEX II. High resolution mass spectra (HRMS) were measured on a Bruker maXis 4G QTOF ESI mass spectrometer. Reversed-phase preparative HPLC was performed on a varian prepstar solvent delivery module equipped with a Gemini NX5u RP18e 250*21.1 mm column (Phenomenex) protected by SecurityGuard PREP Cartridge (Gemini C18 15 × 21.2 mm). Cyclopropanation catalysis samples were analyzed on an Agilent 6890 Series GC System with Agilent CP-Chirasil-Dex CB column (25 m × 0.25 mm × 0.25 μm). The results of C-H insertion catalysis were analyzed by Agilent HP-1 column (30 m × 0.32 mm × 0.25 μm). High-performance liquid chromatography was performed on Agilent 1100 Series with UV-Vis detection.

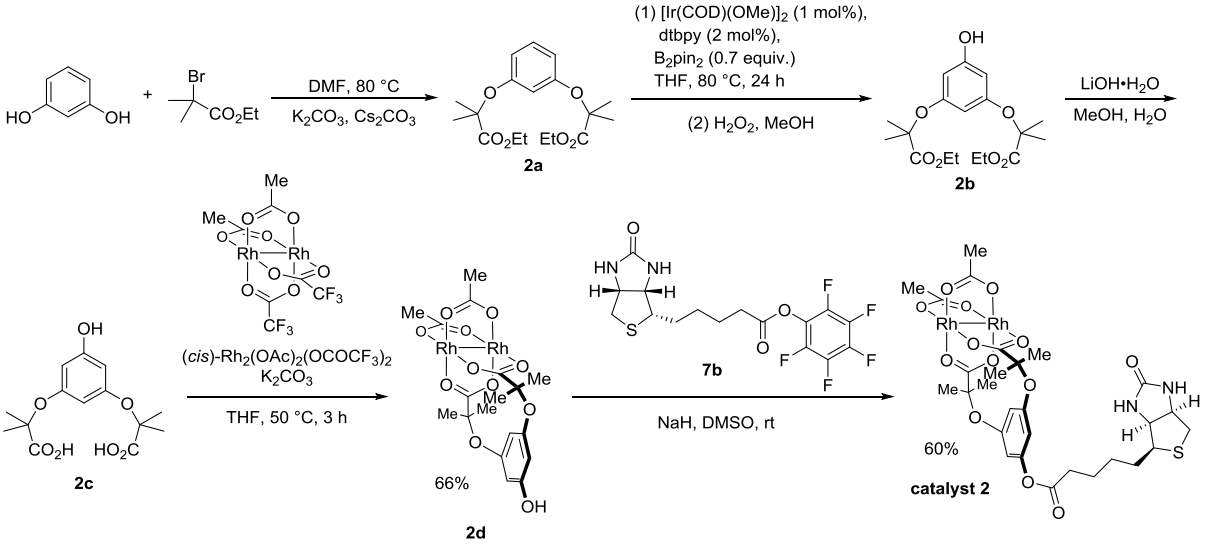
Catalysts synthesis

Synthesis of catalyst 1



Compounds **1a**, **1b**, **1c** and **1** were prepared according to reported procedures.[12]

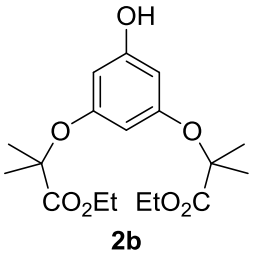
Synthesis of catalyst 2



Compound **2a** was synthesized according to a reported method.[36]

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.07 (t, *J* = 8.2 Hz, 1H), 6.49 (dd, *J* = 8.2, 2.3 Hz, 2H), 6.39 (d, *J* = 2.3 Hz, 1H), 4.24 (q, *J* = 7.1 Hz, 4H), 1.56 (s, 12H), 1.25 (t, *J* = 7.1 Hz, 6H).

diethyl 2,2'-((5-hydroxy-1,3-phenylene)bis(oxy))bis(2-methylpropanoate) 2b



2b

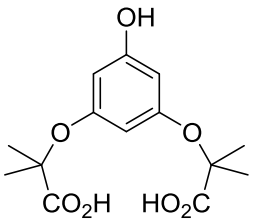
To a 50 mL two-necked round-bottom flask, compound **2a** (0.8957g, 2.65 mmol, 1 equiv.), $[Ir(COD)(OMe)]_2$ (17.6 mg, 26.5 µmol, 1 mol%), 4,4'-di-tert-butylbipyridine (14.2 mg, 52.9 µmol, 2 mol%) and B_2pin_2 (470.5 mg, 1.85 mmol, 0.7 equiv.) were added. The flask was evacuated and charged with N_2 three times. Under a mild N_2-flow, THF (20 mL) was added and the reaction mixture was heated at 80 °C for 48 h. The solvent was removed under reduced pressure, and the resulting crude mixture was purified by flash chromatography (EtOAc / cyclohexane = 1 / 5) to yield a crude product. This product was dissolved in 40 mL MeOH and an H_2O_2 solution (30% in water, 3 mL) was added. The reaction mixture was stirred at rt for 2 h and the solvent was evaporated. Purification by flash chromatography (EtOAc / cyclohexane = 1 / 3) afforded phenol **2b** (0.5795 g, 62%) as colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.03 (d, *J* = 2.1 Hz, 2H), 5.94 (t, *J* = 2.1 Hz, 1H), 5.24 (s, 1H), 4.24 (q, *J* = 7.1 Hz, 4H), 1.56 (s, 14H), 1.25 (t, *J* = 7.1 Hz, 7H).

^{13}C NMR (100 Hz, $CDCl_3$) δ 174.5, 157.0, 156.8, 103.0, 100.9, 79.4, 61.7, 25.5, 14.2.

HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{18}H_{26}NaO_7$, 377.1576; found: 377.1576.

2,2'-((5-hydroxy-1,3-phenylene)bis(oxy))bis(2-methylpropanoic acid) 2c



2c

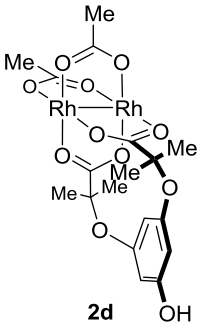
Phenol **2b** (0.5795 g, 1.64 mmol, 1 equiv.) and $LiOH{\cdot}H_2O$ (0.686 g, 16.4 mmol, 10 equiv.) were dissolved in MeOH / H_2O = 4 / 1 (20 mL). The mixture was heated at 60 °C for 3 h. After cooling to rt, the mixture was acidified with 1 M HCl to pH 1 and extracted with EtOAc (30 mL × 3). The organic layer was washed with brine and dried over Na_2SO_4. Concentration under reduced pressure yielded diacid **2c** (0.478 g, 98%) as white solid.

^{1}H NMR (400 MHz, CD_3OD) δ 5.93 (d, *J* = 2.1 Hz, 2H), 5.85-5.84 (m, 1H), 1.44 (s, 12H).

^{13}C NMR (100 Hz, CD_3OD) δ 177.7, 159.5, 158.1, 103.2, 102.0, 80.1, 25.8.

HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{14}H_{18}NaO_7$, 321.0950; found: 321.0950.

Dirhodium complex 2d



2d

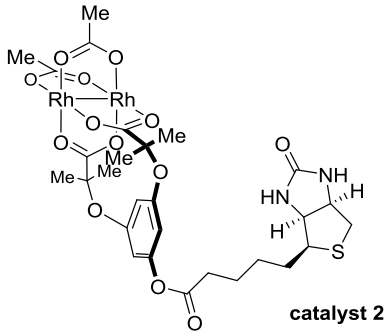
A round-bottom flask was charged diacid **2c** (26.9 mg, 0.09 mmol, 1 equiv.), (*cis*)-$Rh_2(OAc)_2(OCOCF_3)_2$ (49.5 mg, 0.09 mmol, 1 equiv.) and K_2CO_3 (24.9 mg, 0.18 mmol, 2 equiv.). THF (10 mL) was added and the mixture was heated to 50 °C for 3 h. The solvent was removed and crude mixture was purified by flash chromatography (EtOAc / cyclohexane = 1 / 1). Complex **2d** (37.1 mg, 66%) was isolated as a green solid.

^{1}H NMR (400 MHz, CD_3OD) δ 5.92 (m, 2H), 5.62-5.61 (m, 1H), 1.44 (s, 12H), 1.81 (s, 6H), 1.33 (m, 12H).

^{13}C NMR (100 Hz, CD_3OD) δ 193.6, 192.1, 159.3, 157.9, 103.4, 102.2, 81.4, 68.2, 25.5, 23.0.

HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{18}H_{22}NaO_{11}Rh_2$, 642.9170; found: 642.9175.

Catalyst 2



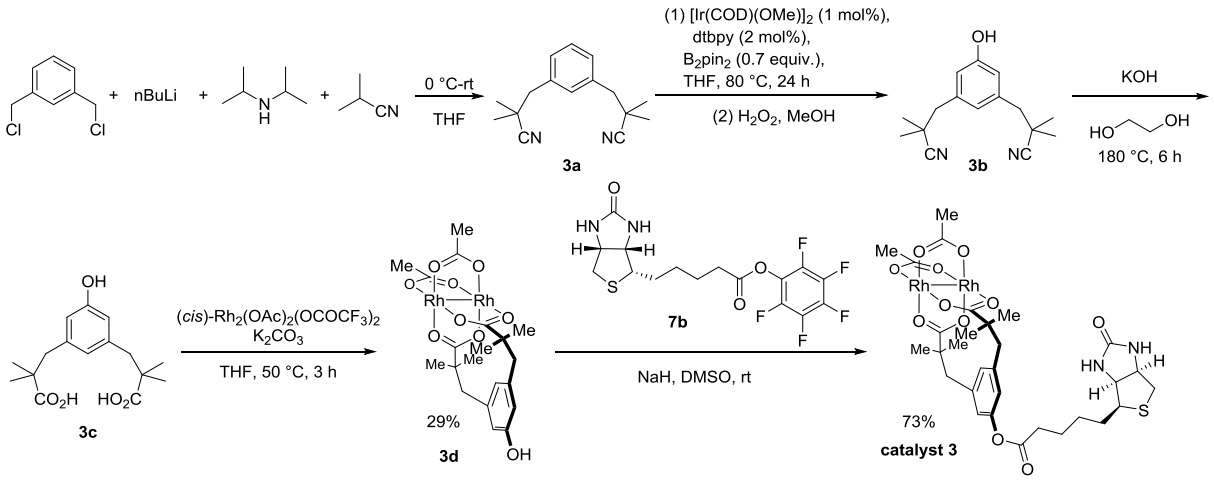
Complex **2d** (14.8 mg, 23.8 µmol, 1 equiv.) and NaH (60%, 0.9 mg, 23.8 µmol, 1 equiv.) were added into a round-bottom flask. DMSO (0.8 mL) was added and the mixture stirred for 1 h at 50 °C. Biotin-pentafluorophenyl ester **7b** (10.7 mg, 26.2 µmol, 1.1 equiv.) was added and the resulting solution was stirred at rt for 3 h. The crude reaction mixture was purified by preparative reversed-phase HPLC to afford catalyst **2** (10.3 mg, 51%) as light purple solid. HPLC gradient: MeCN / H_2O, 2% to 70% for 40 min. t_R = 30 min.

^{1}H NMR (600 MHz, DMSO-d_6) δ 6.46 (s, 1H), 6.37 (s, 1H), 6.27 (d, *J* = 2.0 Hz, 1H), 5.58 (t, *J* = 2.0 Hz, 1H), 4.32 (s, 1H), 4.16 (s, 1H), 3.15 (s, 1H), 2.85-2.84 (m, 1H), 2.61-2.59 (m, 1H), 2.52-2.51 (m, 2H), 1.81 (s, 6H), 1.70-1.62 (m, 3H), 1.52 (s, 1H), 1.45-1.39 (m, 2H), 1.31 (s, 12H).

^{13}C NMR (150 Hz, DMSO-d_6) δ 192.7, 192.2, 171.9, 163.2, 156.1, 151.8, 108.4, 105.5, 80.9, 61.5, 59.7, 55.8, 33.7, 28.4, 28.3, 25.1, 25.1, 24.8, 24.0.

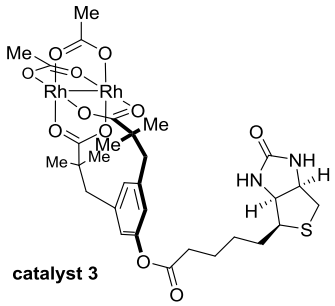
HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{28}H_{36}N_2NaO_{13}Rh_2S$, 868.9946; found: 868.9945.

Synthesis of catalyst 3



Compounds **3a**, **3b**, **3c**, and **3d** were prepared using the same procedures as described in reference 37. Catalyst **3** was synthesized using the same procedures as described for catalyst **2**.

Catalyst 3

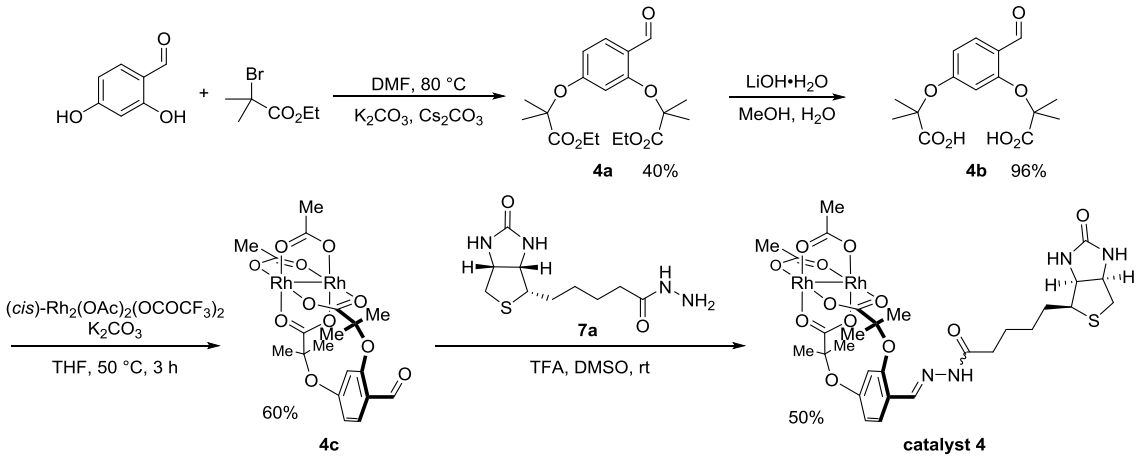


^{1}H NMR (600 MHz, DMSO-d_6) δ 6.59 (s, 2H), 6.47-6.45 (m, 2H), 6.37 (s, 1H), 4.32 (s, 1H), 4.16 (s, 1H), 3.15 (s, 1H), 2.85-2.84 (m, 1H), 2.61-2.59 (m, 1H), 2.54 (s, 6H), 1.78 (s, 6H), 1.67-1.61 (m, 3H), 1.52 (s, 1H), 1.41-1.40 (m, 2H), 0.91 (s, 12H).

^{13}C NMR (150 Hz, DMSO-d_6) δ 196.7, 191.0, 171.6, 162.7, 149.1, 138.8, 127.7, 120.8, 61.0, 59.1, 55.3, 45.8, 45.8, 40.4, 33.3, 27.9, 25.2, 24.3, 23.5.

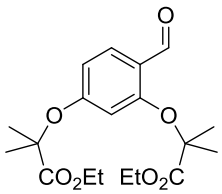
HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{30}H_{40}N_2NaO_{11}Rh_2S$, 865.0361; found: 865.0354.

Synthesis of catalyst 4



Compound **4a**, **4b**, **4c** and catalyst **4** were prepared using the same procedures as described for compound **1a**, **1b**, **2c**, **1**.

diethyl 2,2'-((4-formyl-1,3-phenylene)bis(oxy))bis(2-methylpropanoate) 4a



4a

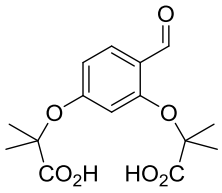
2,4-dihydroxybenzaldehyde (0.691 g, 5 mmol, 1.0 equiv.) was added to a flame-dried round bottom flask charged with a magnetic stir bar, along with K_2CO_3 (3.32 g, 24 mmol, 4.8 equiv.) and KI (0.83 g, 5 mmol, 1.0 equiv.). The solid reagents were then dissolved in 20 mL dry DMF under N_2. Ethyl 2-bromo-2-methylpropanoate (2.93 mL, 20 mmol, 4.0 equiv.) was then added dropwise to the reaction mixture. The reaction was heated to 60 °C overnight. After cooling to rt, the mixture was quenched with a saturated NH_4Cl solution and diluted with 20 mL H_2O and 20 mL Et_2O. The aqueous layer was extracted twice with Et_2O (20 mL × 2). The organic layer was collected, dried with Na_2SO_4 and concentrated. The crude product was purified by flash chromatography (EtOAc / cyclohexane = 1 / 5 ~ 1 / 4) to afford the product **4a** (0.73 g, 40%) as a colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.33 (s, 1H), 7.75 (d, *J* = 8.7 Hz, 1H), 6.46 (dd, *J* = 8.7, 2.1 Hz, 1H), 6.20 (d, *J* = 2.2 Hz, 1H), 4.29-4.20 (m, 4H), 1.66 (s, 6H), 1.62 (s, 6H), 1.27-1.21 (m, 6H).

^{13}C NMR (100 Hz, $CDCl_3$) δ 188.9, 173.5, 161.7, 159.8, 129.7, 121.7, 111.7, 106.6, 80.4, 79.7, 62.0, 61.9, 25.6, 25.6, 14.1.

HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{19}H_{26}NaO_7$, 389.1576; found: 389.1576.

2,2'-((4-formyl-1,3-phenylene)bis(oxy))bis(2-methylpropanoic acid) 4b



4b

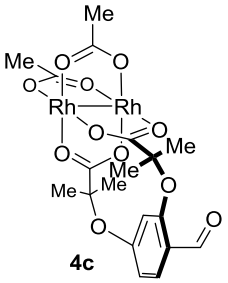
Diester **4a** (0.159 g, 0.434 mmol, 1 equiv.) and LiOH · H_2O (0.182 g, 4.34 mmol, 10 equiv.) were dissolved in MeOH / H_2O = 4 / 1 (5 mL). The mixture was heated at 60 °C for 3 h. After cooling to rt, the mixture was acidified with 1 M HCl solution to pH 1 and extracted with EtOAc (20 mL × 3). The organic layer was washed with brine and dried over Na_2SO_4. Concentration under reduced pressure yielded diacid **4b** (0.13 g, 97%) as light yellow solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.36 (br, 2H), 10.19 (br, 1H), 7.72 (d, *J* = 8.8 Hz, 1H), 6.56 (ddd, *J* = 8.6, 2.2, 0.7 Hz, 1H), 5.94 (d, *J* = 2.2 Hz, 1H), 1.66 (s, 6H), 1.63 (s, 6H).

^{13}C NMR (100 Hz, $CDCl_3$) δ 188.7, 177.8, 177.7, 160.9, 158.5, 131.1, 120.4, 113.0, 103.5, 79.4, 78.9, 25.4, 25.3.

HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{15}H_{18}NaO_7$, 333.0950; found: 333.0942.

Dirhodium complex 4c



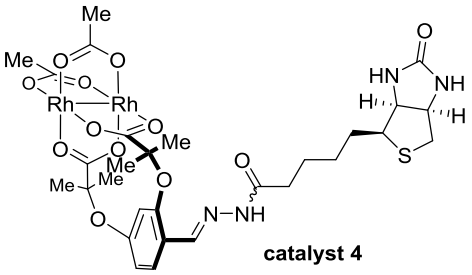
A round bottom flask was charged phenol diacid **4b** (30.7 mg, 0.099 mmol, 1.1 equiv.), (*cis*)-$Rh_2(OAc)_2(OCOCF_3)_2$ (49.5 mg, 0.09 mmol, 1 equiv.) and K_2CO_3 (24.9 mg, 0.18 mmol, 2 equiv.). THF (10 mL) was added and the mixture was heated to 50 °C for 3 h. The solvent was removed under reduced pressure and the crude mixture was purified by flash chromatography (EtOAc / cyclohexane = 1 / 1) to afford complex **4c** (44 mg, 77%) as a green solid.

^{1}H NMR (400 MHz, CD_3OD) δ 10.65 (s, 1H), 7.62 (d, *J* = 8.8 Hz, 1H), 6.47 (ddd, *J* = 8.7, 2.3, 0.9 Hz, 1H), 5.65 (d, *J* = 2.2 Hz, 1H), 1.82 (s, 6H), 1.46-1.43 (m, 12H).

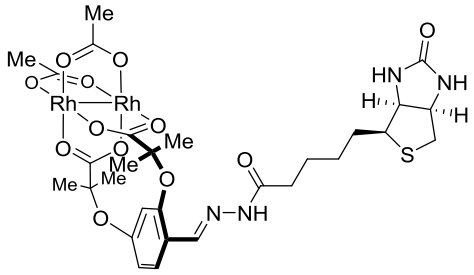
^{13}C NMR (100 Hz, CD_3OD) δ 192.8, 192.7, 192.4, 190.1, 163.2, 161.2, 129.8, 121.5, 114.2, 106.5, 82.3, 81.6, 61.9, 52.0, 31.3, 28.9, 28.0, 25.7, 23.0.

HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{19}H_{22}NaO_{11}Rh_2$, 654.9170; found: 654.9162.

Catalyst 4



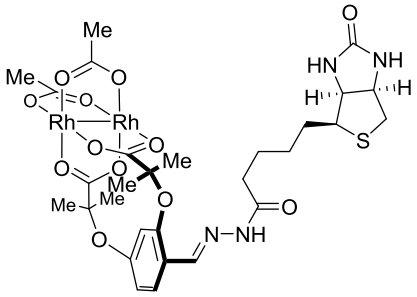
To a solution of complex **4c** (10.8 mg, 17 µmol, 1.0 equiv.) and biotin hydrazide **7a** (4.4 mg, 17 µmol, 1.0 equiv.) in dry DMSO (1 mL), TFA (1.3 µL, 17 µmol, 1.0 equiv.) was added. The reaction mixture was stirred for 3h at rt. After dilution with MeCN / H_2O = 1 / 2 (3 mL), the crude reaction mixture was purified by preparative HPLC to afford catalyst **4** (10 mg, 67%) as light purple solid. HPLC gradient: MeCN / H_2O, 10% to 70% for 40 min. t_R = 21 min.



One isomer:

^{1}H NMR (600 MHz, DMSO-d_6) δ 11.22 (s, 1H), 8.28 (s, 1H), 7.66 (d, J = 4.0 Hz, 1H), 6.56-6.55 (m, 1H), 6.44 (s, 1H), 6.35 (s, 1H), 5.68 (d, J = 2.3 Hz, 1H), 4.31-4.30 (m, 1H), 4.15-4.13 (m, 1H), 3.13 (m, 1H), 2.84-2.83 (m, 1H), 2.59 (s, 1H), 2.16 (t, J = 7.4 Hz, 2H), 1.81 (s, 6H), 1.68-1.56 (m, 4H), 1.52-1.46 (m, 2H), 1.34-1.32 (m, 12H).

^{13}C NMR (150 Hz, DMSO-d_6) δ 191.7, 192.1, 173.9, 168.2, 162.7, 156.5, 141.0, 126.1, 119.4, 115.2, 107.8, 81.2, 61.0, 59.1, 55.4, 40.0, 34.0, 28.8, 28.3, 25.0, 24.6, 23.5.



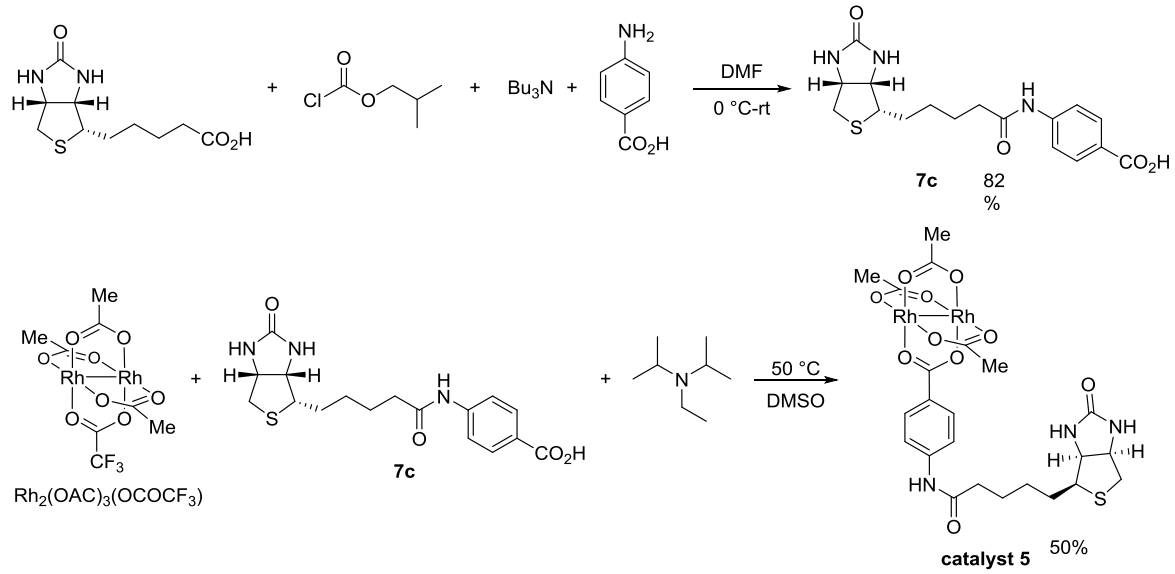
Other isomer:

^{1}H NMR (600 MHz, DMSO-d_6) δ 11.04 (s, 1H), 8.15 (s, 1H), 7.64 (d, J = 4.0 Hz, 1H), 6.54-6.52 (m, 1H), 6.43 (s, 1H), 6.35 (s, 1H), 5.64 (d, J = 2.3 Hz, 1H), 4.31-4.30 (m, 1H), 4.15-4.13 (m, 1H), 3.13 (m, 1H), 2.82-2.81 (m, 1H), 2.57 (s, 1H), 2.16 (t, J = 7.4 Hz, 2H), 1.81 (s, 6H), 1.68-1.56 (m, 4H), 1.52-1.46 (m, 2H), 1.34-1.32 (m, 12H).

^{13}C NMR (150 Hz, DMSO-d_6) δ 191.7, 192.2, 174.5, 173.9, 162.7, 156.6, 138.4, 125.8, 118.9, 115.2, 107.3, 80.9, 61.0, 59.2, 55.4, 40.0, 34.0, 28.8, 28.3, 25.0, 24.6, 23.5.

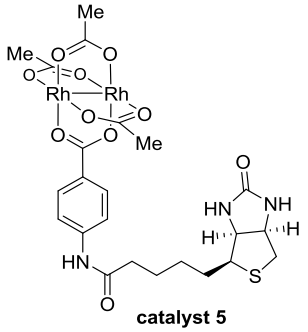
HRMS (ESI, pos.) m/z: $[M + Na]^+$ calcd for $C_{29}H_{38}N_4NaO_{12}Rh_2S$, 895.0215; found: 895.0211.

Synthesis of catalyst 5



Compound **7c** was prepared according to the reported procedure.[38]

catalyst 5



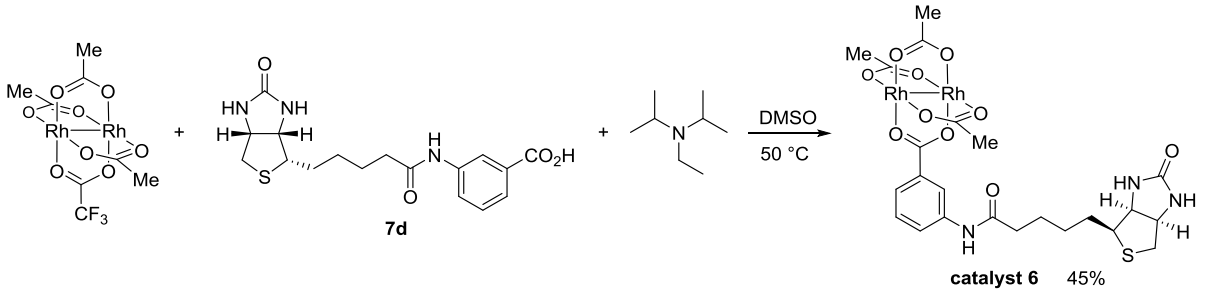
$Rh_2(OAc)_3(OCOCF_3)$ (7.6 mg, 15.3 μmol, 1 equiv.), compound **7c** (5.6 mg, 15.3 μmol, 1 equiv.) and N, N-Diisopropylethylamine (3.2 μL, 18.4 μmol, 1.2 euquiv.) were added into a round bottom flask. Dry DMSO (1 mL) was charged and the mixture was heated to 50 °C for 2 h. The reaction mixture was subjected to preparative HPLC for purification. Final product **5** (5.7 mg, 50%) was obtained as light purple solid. HPLC gradient: MeCN / H_2O, 2% to 60% for 40 min. t_R = 19 min.

^{1}H NMR (600 MHz, DMSO-d_6) δ 10.12 (s, 1H), 7.71 (d, J = 8.8 Hz, 2H), 7.59 (d, J = 8.8 Hz, 2H), 6.42 (s, 1H), 6.35 (s, 1H), 4.31-4.29 (s, 1H), 4.13 (s, 1H), 3.12-3.11 (m, 1H), 2.83-2.80 (m, 1H), 2.62-2.56 (m, 1H), 2.31 (t, J = 7.0 Hz, 2H), 1.84 (s, 3H), 1.77 (m, 6H), 1.67-1.55 (m, 3H), 1.50-1.48 (s, 1H), 1.40-1.30 (m, 2H).

^{13}C NMR (150 Hz, DMSO-d_6) δ 191.4, 191.3, 184.3, 171.6, 162.7, 143.0, 129.5, 125.4, 117.8, 61.0, 59.2, 55.4, 40.4, 36.3, 28.2, 28.0, 24.9, 23.7, 23.6.

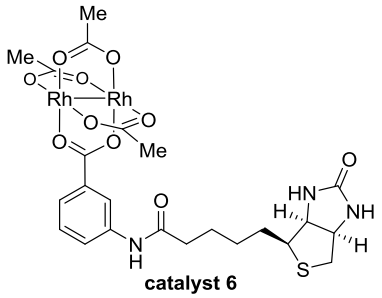
HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{23}H_{29}N_3NaO_{10}Rh_2S$, 767.9581; found: 767.9577.

Synthesis of catalyst 6



Compound **7d** and **6** were synthesized in the same way using **7c** and **5** respectively.

Catalyst 6



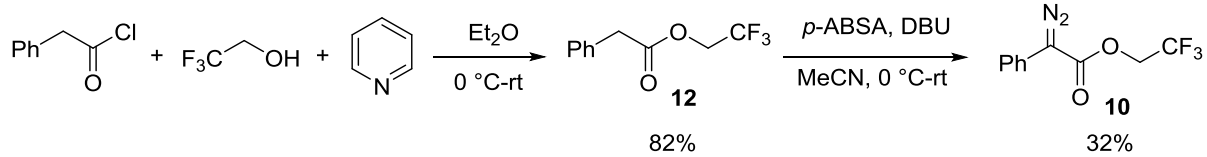
$Rh_2(OAc)_3(OCOCF_3)$ (9.9 mg, 20 μmol, 1 equiv.), compound **7d** (8.0 mg, 22 μmol, 1.1 equiv.) and N, N-Diisopropylethylamine (4.2 μL, 24 μmol, 1.2 equiv.) was charged into a round bottom flask. DMSO (1 mL) was charged and heated to 50 °C for 2 h. The reaction mixture was purified by preparative HPLC. The pure product **6** (7.5 mg, 50%) was obtained as light purple solid. HPLC gradient: MeCN / H_2O, 2% to 50% for 40 min. t_R = 25 min.

^{1}H NMR (600 MHz, DMSO-d_6) δ 9.99 (s, 1H), 8.09 (s, 1H), 7.70 (d, *J* = 7.8 Hz, 1H), 7.42 (d, *J* = 7.8 Hz, 1H), 7.30 (t, *J* = 7.9 Hz, 1H), 6.44 (s, 1H), 6.37 (s, 1H), 4.32 (s, 1H), 4.15 (s, 1H), 3.14-3.08 (m, 1H), 2.84-2.83 (m, 1H), 2.60-2.58 (m, 1H), 2.30 (t, *J* = 7.0 Hz, 2H), 1.84 (s, 3H), 1.78 (m, 6H), 1.67-1.57 (m, 3H), 1.52 (s, 1H), 1.42-1.32 (m, 2H).

^{13}C NMR (150 Hz, DMSO-d_6) δ 191.5, 191.3, 184.4, 171.4, 162.7, 139.1, 131.3, 128.5, 122.8, 122.6, 118.8, 61.0, 59.2, 55.4, 40.3, 36.3, 28.1, 28.0, 25.1, 23.6, 23.6.

HRMS (ESI, pos.) *m/z*: $[M + Na]^+$ calcd for $C_{23}H_{29}N_3NaO_{10}Rh_2S$, 767.9581; found: 767.9584.

Synthesis of the C-H insertion substrate 10 and product 11

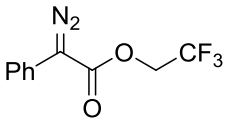


2, 2, 2-trifluoroethyl 2-phenylacetate 12

2,2,2-Trifluoroethanol (0.57 mL, 7.5 mmol, 1.5 equiv.) and pyridine (0.8 mL, 10 mmol, 2 equiv.) were dissolved in 25 mL dry Et_2O and stirred for 20 min. Phenylacetyl chloride (0.66 mL, 5 mmol, 1 equiv.) was slowly added at 0 °C and the reaction mixture was stirred overnight at rt. The reaction was quenched with water and extracted with Et_2O (30 mL × 3). The crude mixture was concentrated and purified by flash column chromatography (EtOAc / cyclohexane = 1 / 10 ~ 1 / 8) to yield the product (82%) as white solid. Spectral data for **12** was identical to those previously reported.[39]

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33-7.29 (m, 5H), 4.47 (q, *J* = 8.4 Hz, 2H), 3.73 (s, 2H).

2, 2, 2-trifluoroethyl 2-diazo-2-phenylacetate 10



10

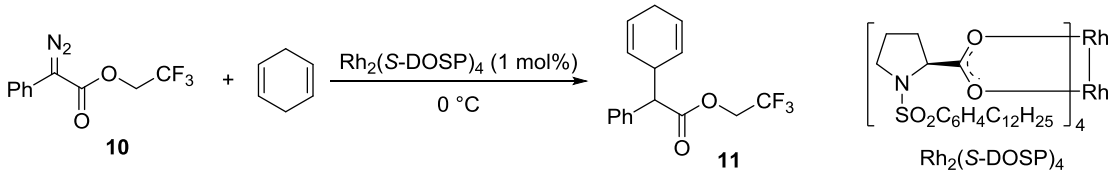
Compound **13** (0.2367 g, 1.1 mmol, 1.0 equiv.) was dissolved in acetonitrile (20 mL) together with *p*-ABSA (0.39 g, 1.63 mmol, 1.5 equiv.) and the solution cooled to 0 °C. Then, DBU (0.32 mL, 2.17 mmol, 2.0 equiv.) was added drop-wise. The solution was stirred for 6 hours at rt and quenched with saturated aqueous NH_4Cl (15 mL) and H_2O (20 mL). The solution was extracted with Et_2O (20 mL × 3) and the organic layer separated, dried over Na_2SO_4, and concentrated. The crude material was purified by column chromatography (Et_2O / cyclohexane = 1 / 20) to give the diazo as a yellow solid (85.2 mg, 32% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.48-7.39 (m, 4H), 7.25-7.21 (m, 1H), 4.65 (q, *J* = 8.3 Hz, 2H).

^{13}C NMR (100 Hz, $CDCl_3$) δ 129.1, 126.4, 124.5, 124.3, 124.1, 121.5, 60.3 (q, *J* = 36.7 Hz).

^{19}F NMR (376.5 MHz, $CDCl_3$) δ -73.5 (t, *J* = 8.4 Hz, 3F).

2, 2, 2-trifluoroethyl 2-(cyclohexa-2, 5-dien-1-yl)-2-phenylacetate 11



Diazo **10** (48.8 mg, 0.2 mmol, 1 equiv.) and 1, 4-cyclohexadiene (97.5 µL, 1 mmol, 5 equiv.) were charged into a 1.5 mL vial and cooled to 0 °C. Rh_2(*S*-DOSP)$_4$ (1.32 mg, 0.002 mmol, 1 mol%) was carefully added to the mixture. The reaction mixture was stirred for 30 min at 0 °C and subjected to column chromatography (Et_2O / cyclohexane = 1 / 20) for purification. The product **11** was obtained (78% yield) as colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37-7.28 (m, 5H), 5.86-5.81 (m, 1H), 5.73-5.68 (m, 2H), 4.62-4.52 (m, 1H), 4.41-4.31 (m, 1H), 3.35-3.46 (m, 2H), 2.64-2.60 (m, 2H).

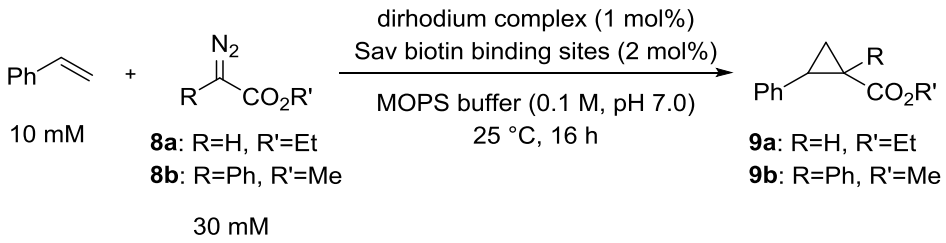
^{13}C NMR (100 Hz, $CDCl_3$) δ 171.4, 135.7, 128.6, 127.7, 126.8, 126.3, 125.8, 125.3, 60.3 (q, *J* = 36.7 Hz), 57.9, 38.5, 26.4.

^{19}F NMR (376.5 MHz, $CDCl_3$) δ -73.0 (t, *J* = 8.4 Hz, 3F).

HRMS (ESI, pos.) *m/z*: $[M+Na]^+$ calcd for $C_{16}H_{15}F_3NaO_2$, 319.0922; found: 319.0920.

General procedure of cyclopropanation (styrene and 8a) and C-H insertion (10 and 1, 4-cyclohexadiene) catalyzed by the artificial carbene transferase

cyclopropanation (styrene with 8a or 8b)

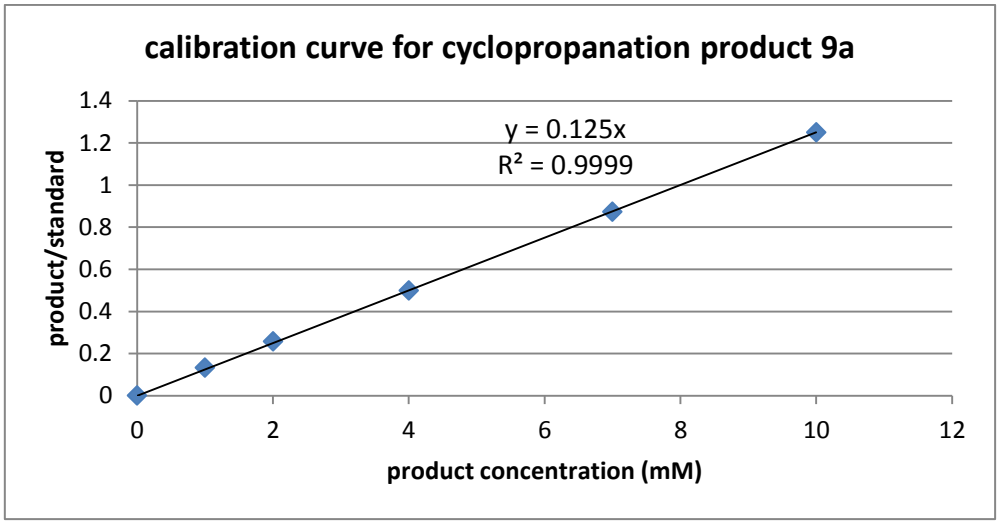


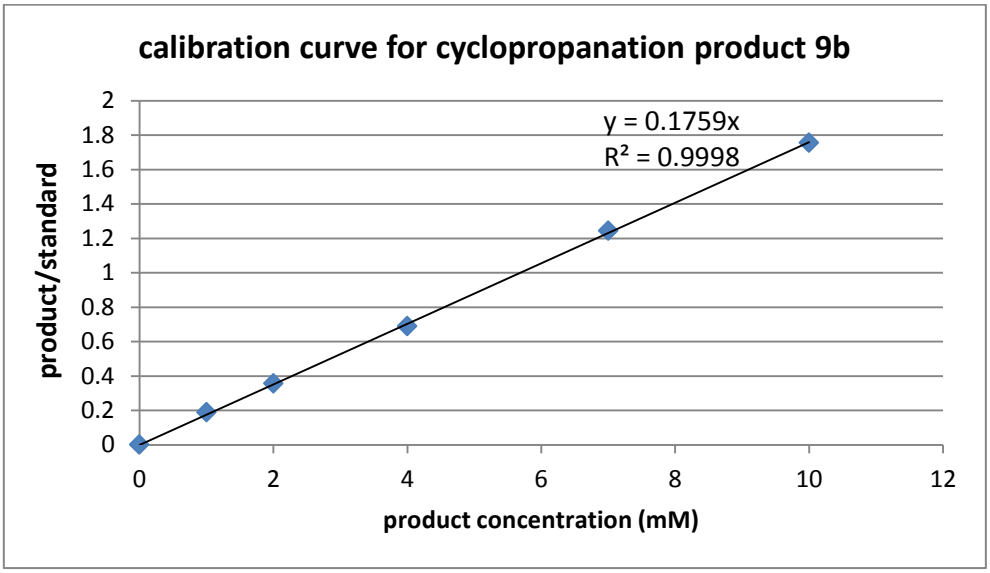
A series of stock solutions were prepared: biotinylated dirhodium catalyst (4 mM in DMSO), streptavidin stock (biotin binding sites 800 μM in MQ water), styrene stock (0.8 M in DMSO), ethyl diazoacetate **8a** or diazo(phenyl)acetate **8b** stock (2.4 M in DMSO) and MOPS buffer (0.2 M, pH 7.0). In a small glass reaction vial, MOPS buffer stock solution (200 μL), protein stock solution (100 μL), catalyst DMSO stock solution (10 μL) and MQ water (80 μL) were added and incubated at 25 °C for 10 min. Then, the styrene stock solution (5 μL) and the diazo compound stock solution (5 μL) were added, and the reaction mixture was shaked in a ThermoMixer for 16 h at 25 °C.

Upon completion of the reaction, the mixture was quenched with a pyridine solution (10 μL, 1 M in H_2O) and the internal standard 1, 3, 5-Trimethoxybenzene (50 μL, 100 mM in ethyl acetate) was added, followed by EtOAc (950 μL). The mixture was thoroughly shaken, and the upper organic phase was transferred to an eppendorf tube, dried with Na_2SO_4 and centrifuged at 14'000 rpm for 10 minutes. The supernatant (700 μL) was transferred in an HPLC vial and subjected to GC analysis to determine the conversion.

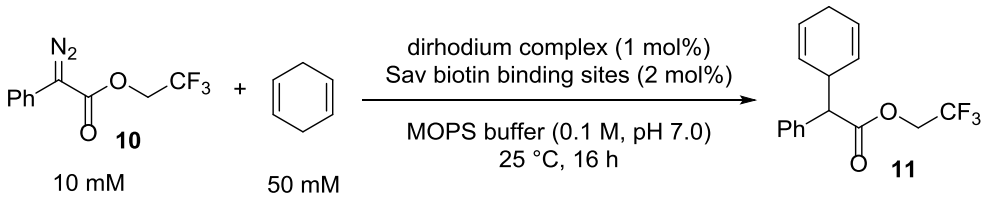
GC analysis. Column: Agilent CP-Chirasil-Dex CB column. Method: $V_{injected}$: 1 μL. Catalysis samples were introduced into the instrument via a split (1:20) injection at a flow rate of 1.7 mL/min helium gas. The temperature program used for product **9a**: 130 °C for 20 minutes. T_R 9.0 min (internal standard 1, 3, 5-Trimethoxybenzene), 11.3 min (product *cis* isomer), 12.3 min (*cis* isomer), 12.8 min (*trans* isomer), 13.3 min (*trans* isomer).

The temperature program used for product **9b**: 170 °C for 30 minutes. T_R 3.4 min (internal standard 1, 3, 5-Trimethoxybenzene), 15.3 min (product *cis* isomer), 16.1 min (*cis* isomer).





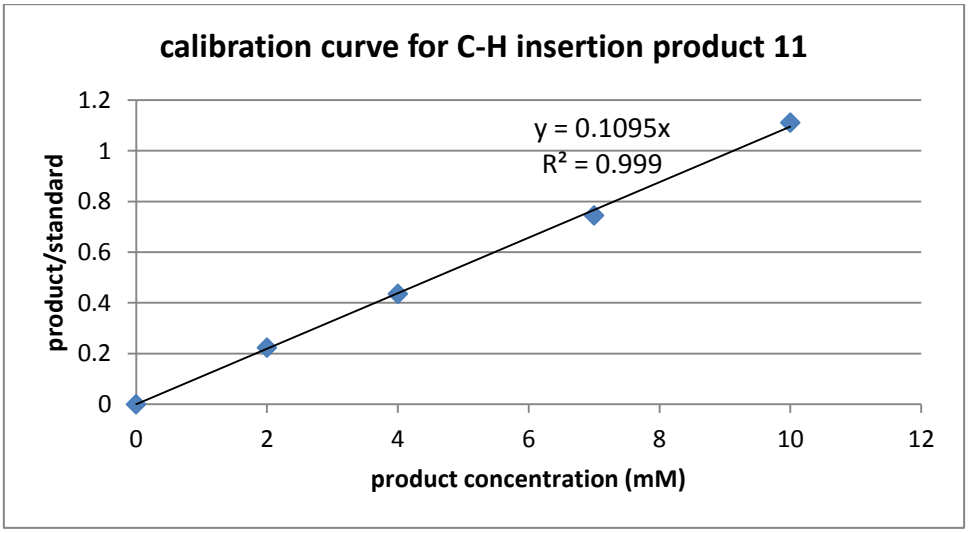
C-H insertion (10 and 1, 4-cyclohexadiene)



In a small glass vial, MOPS buffer stock solution (200 µL), protein stock solution (100 µL), catalyst DMSO stock solution (10 µL) and MQ water (80 µL) were added and incubated at 25 °C for 10 min. Then, the diazo stock solution (5 µL) and 1, 4-cyclohexadiene stock solution (5 µL) were added, and the reaction mixture was shaken in a ThermoMixer for 16 h at 25 °C.

The mixture was then quenched with a pyridine solution (10 µL, 1 M in H_2O) and the internal standard 1, 3, 5-Trimethoxybenzene (50 µL, 100 mM in ethyl acetate) was added, followed by EtOAc (950 µL). The solution mixture was thoroughly mixed and the upper organic layer was transferred to an eppendorf tube, dried with Na_2SO_4 and centrifuged at 14'000 rpm for 10 minutes. The supernatant (700 µL) was transferred in an HPLC vial and subjected to GC analysis to determine the conversion.

GC analysis. Column: Agilent HP-1 column. Method: $V_{injected}$: 1 µL. Catalysis samples were injected via a split (1:20) injection at a flow rate of 1.7 mL/min helium gas. The temperature program for product **11**: 170 °C for 20 minutes. T_R 2.7 min (internal standard 1, 3, 5-Trimethoxybenzene), 4.6 min (product).



ICP-MS analysis for dirhodium uptake in the periplasm screening

In order to quantify the rhodium content of the cells, ICP-MS measurements were performed in duplicate using our reported protocol.[18] Measured and calculated results are shown in table 1.

Table 1 Rhodium uptake determined by ICP-MS upon incubation of catalyst **2** with *E. coli* cells secreting Sav in their periplasm of *E. coli cells.* Based on the [Rh] determined by ICP-MS, the corresponding TON can be determined.

	rhodium added during for incubation (nmol)	amount of Rh determined after incubation and washing (nmol)	Uptake (%)	calculated TON
empty	50	3.2 ± 0.3	6 ± 1	10 ± 3
WT	50	8.3 ± 0.8	17 ± 2	9 ± 0
S112C	50	6.4 ± 0.1	13 ± 0	17 ± 3
S112D	50	6.0 ± 0.5	12 ± 1	11 ± 3

K121C	50	8.0 ± 0.2	16 ± 0	20 ± 2
K121M	50	8.9 ± 0.1	18 ± 0	11 ± 3

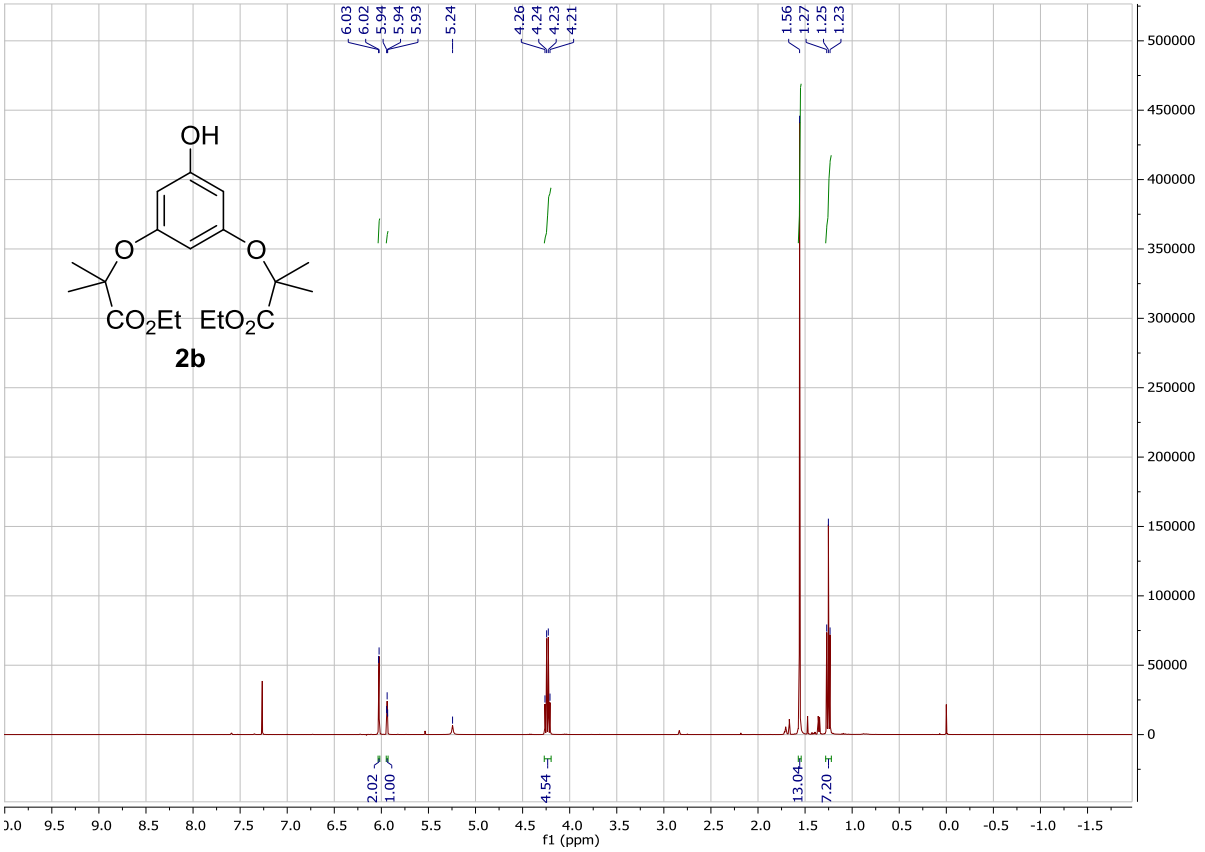
General procedure for cyclopropanation screening in the periplasm of *E. coli*

The TOP10(DE3)_pET30 *E. coli* strain and kanamycin antibiotic were used for the expression of periplasmic OmpA tagged Sav WT, S112C, S112D, K121C and K121M. An overnight culture of the above mutants was inoculated to a main culture (50 mL modified ZYM-5052 rich induction medium[6] in 250 mL erlenmeyer flask) for 3 h shaking (37 °C, 220 rpm). Then, the cultures were induced by IPTG (final concentration 50 µM) and continued shaking for 4 h (25 °C, 220 rpm) to express and secrete the periplasmic Sav.

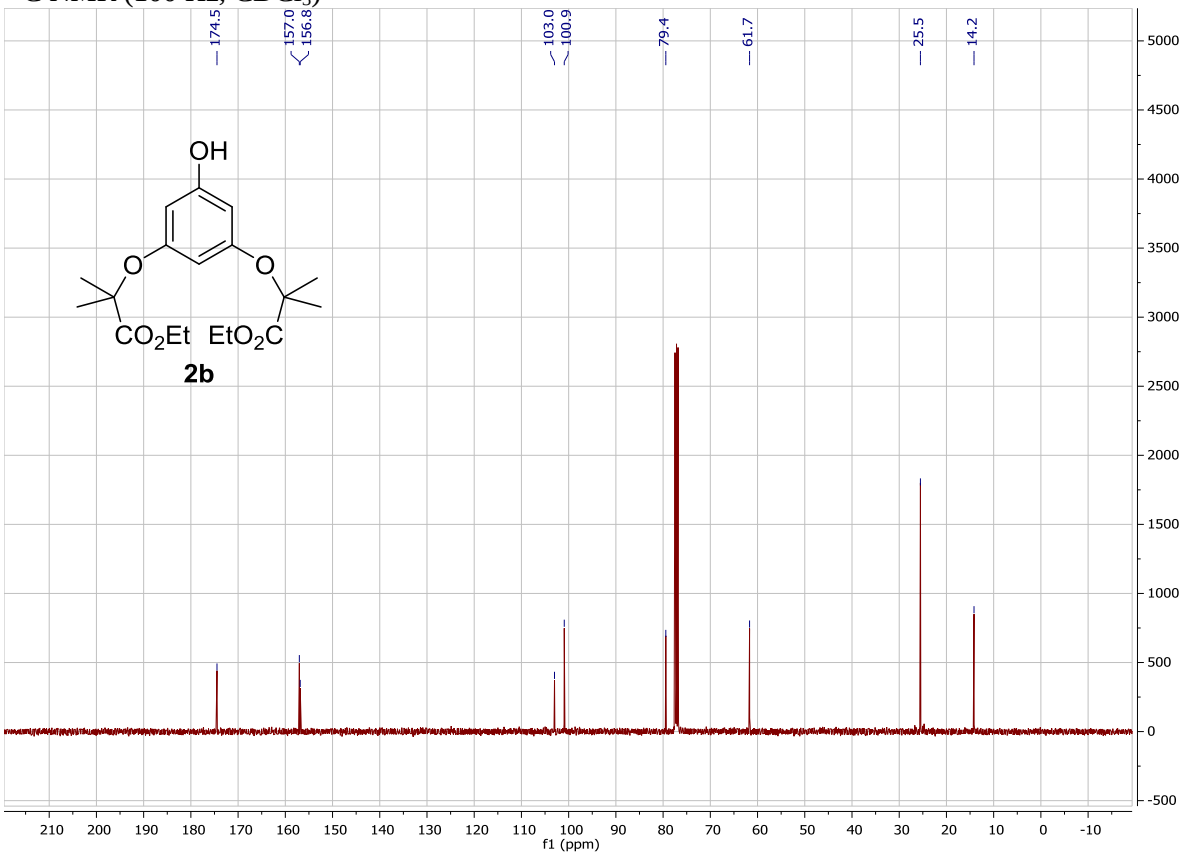
The cells were harvested into a 96-well plate by transferring OD 3/mL cells to each well on the plate. The plate was centrifuged (3200 g, 4 °C, 20 min) and the supernatant was discarded. The cell pellets were washed with MOPS buffer (50 mM, NaCl 0.9%, pH 7.0), spun down and the supernatant was discarded. Then, the cell pellets were incubated in the above MOPS buffer containing additional 50 µM dirhodium catalyst **2** on ice for 0.5 h. The plate was centrifuged and the supernatant was discarded, followed by a second wash with MOPS buffer. The catalysis buffer in MOPS (200 µL of 0.2 M stock solution, pH 7.0) was added into each well. The individual cell pellets were resuspended and transferred to HPLC glass vials. Then styrene and the diazo substrate were added to the glass vials and the catalysis was performed at 25 °C for 16 h. The workup and the GC analysis allowed to quantify the activity of each mutant.

NMR Spectra:

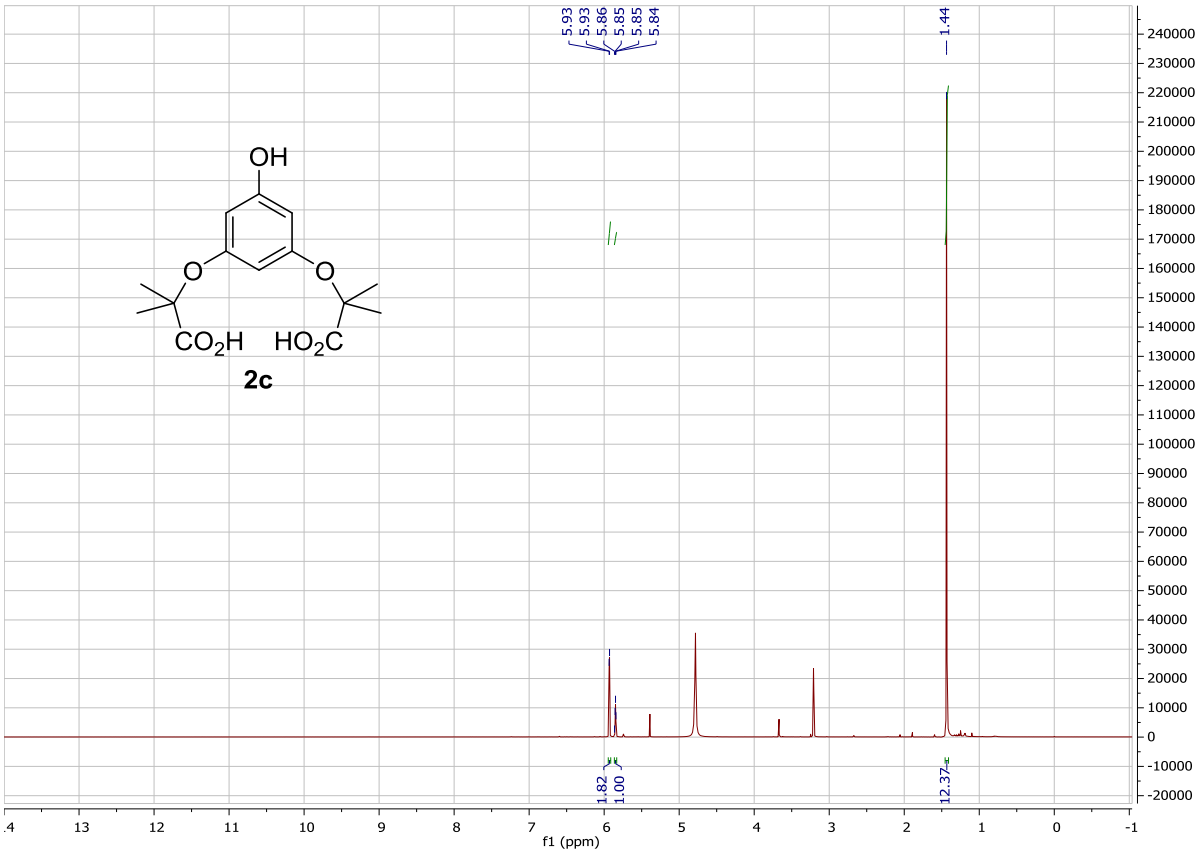
^{1}H NMR (400 MHz, $CDCl_3$)



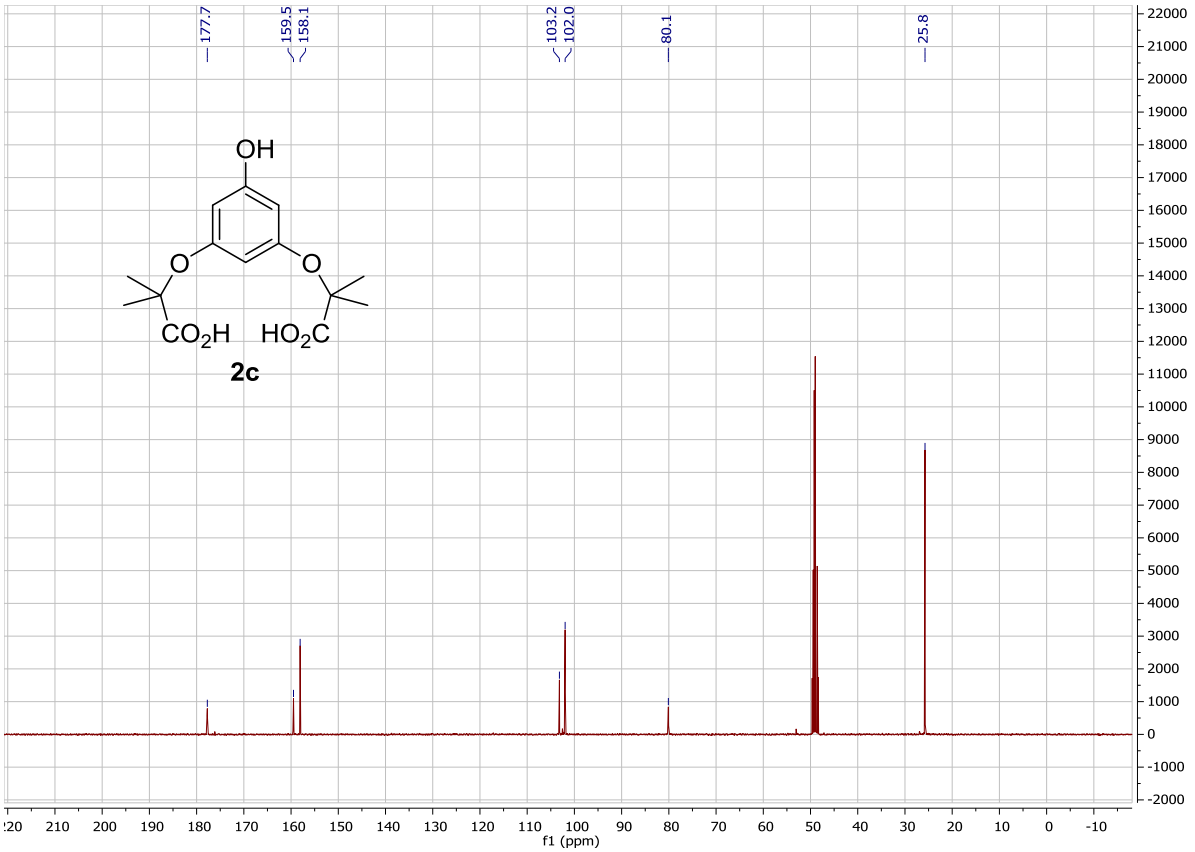
^{13}C NMR (100 Hz, $CDCl_3$)



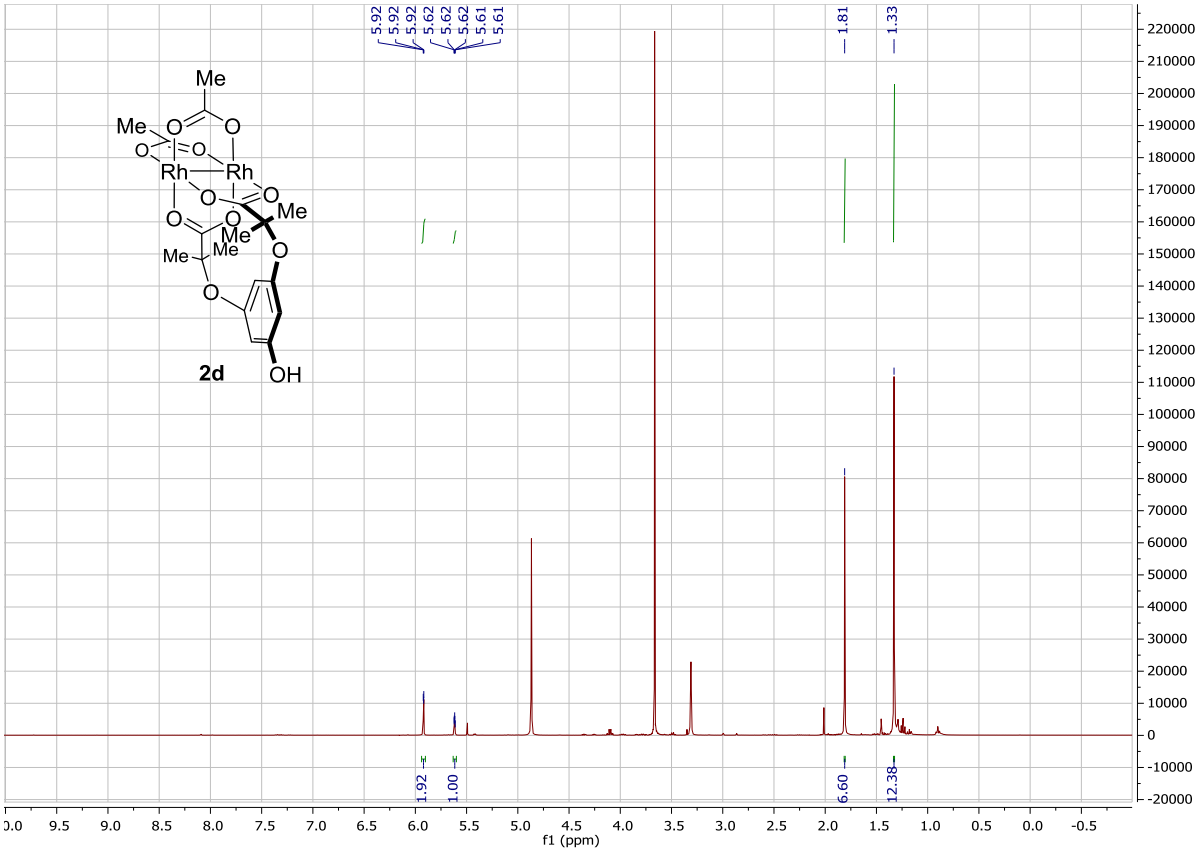
^{1}H NMR (400 MHz, $CDCl_3$)



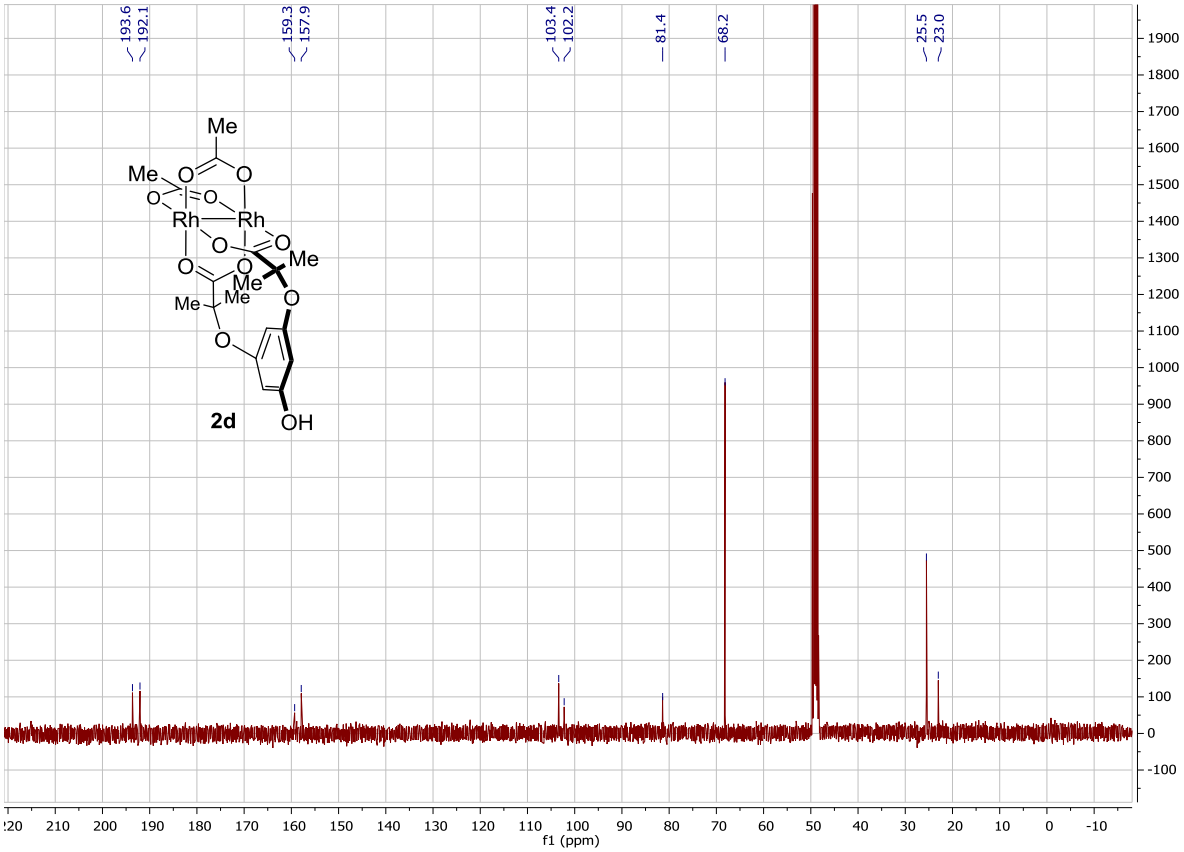
^{13}C NMR (100 Hz, $CDCl_3$)



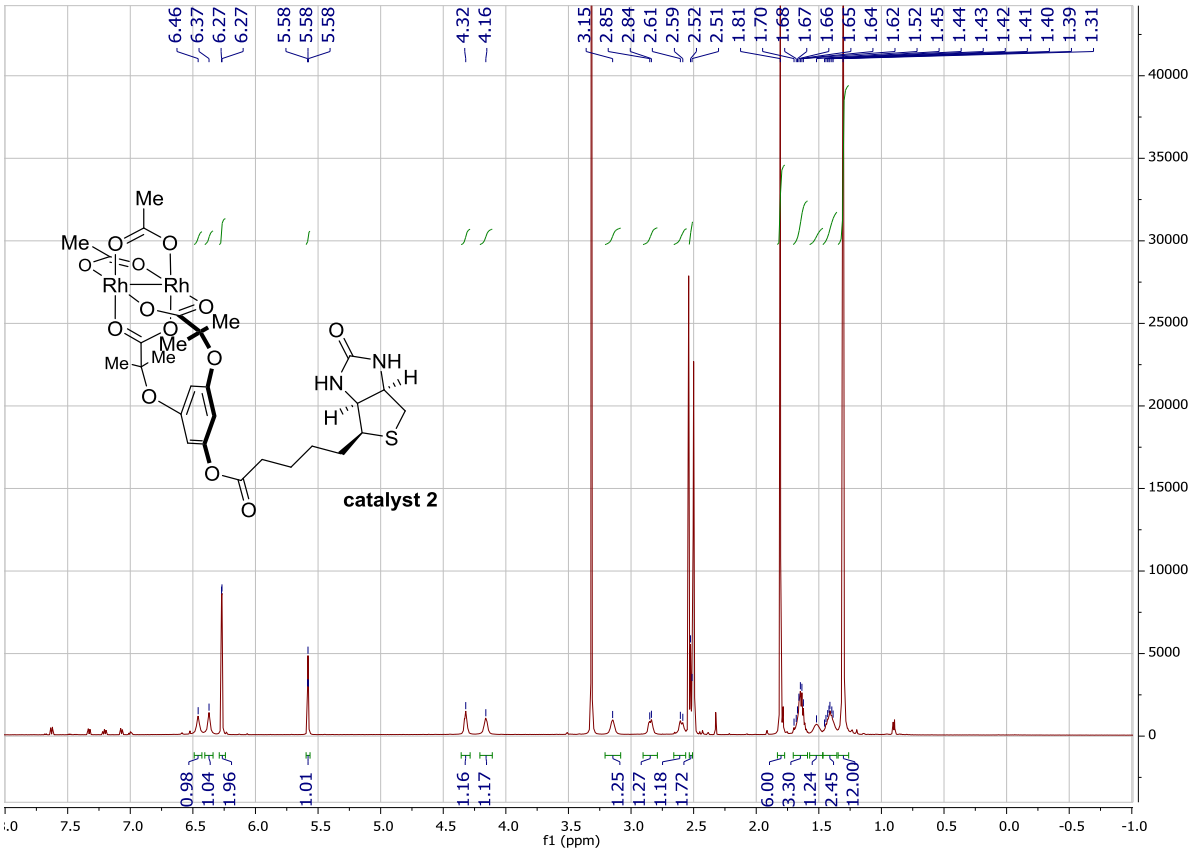
^{1}H NMR (400 MHz, CD_3OD)



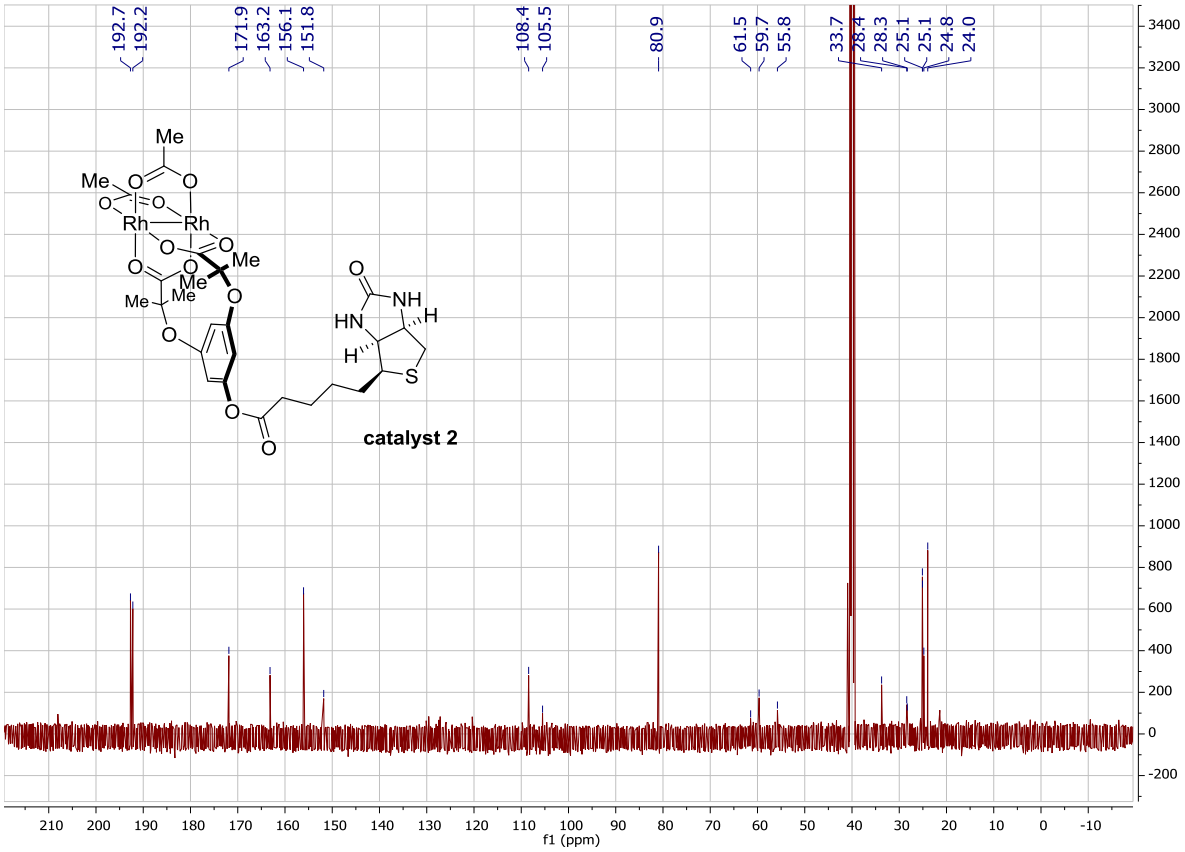
^{13}C NMR (100 Hz, CD_3OD)



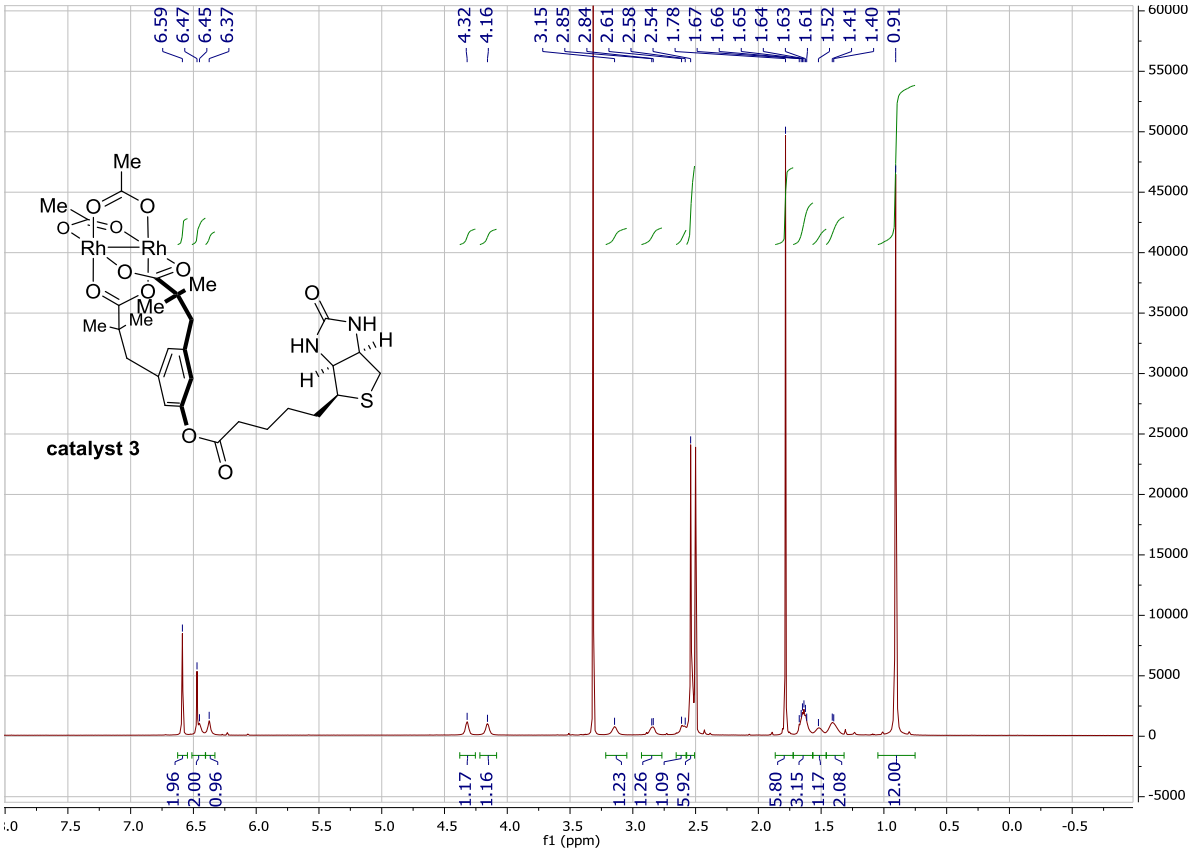
^{1}H NMR (600 MHz, DMSO-d_6)



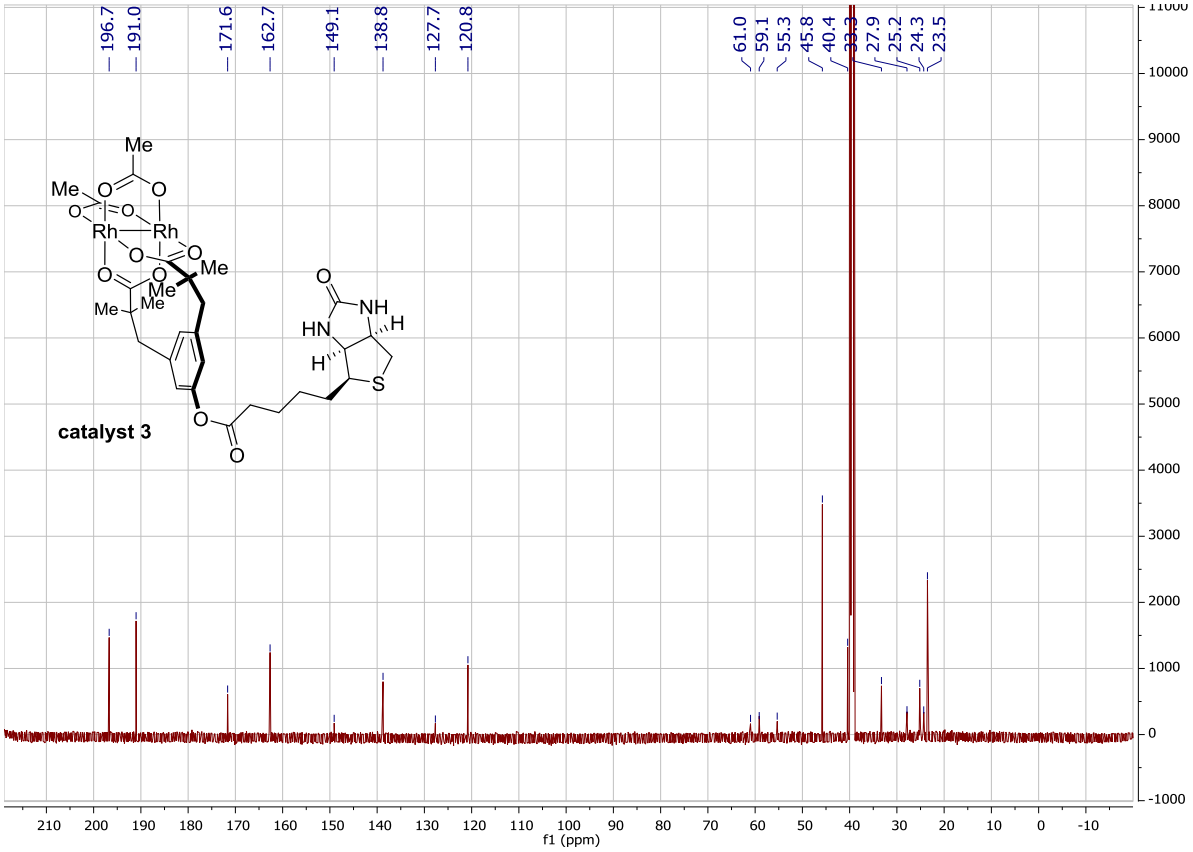
^{13}C NMR (150 Hz, DMSO-d_6)



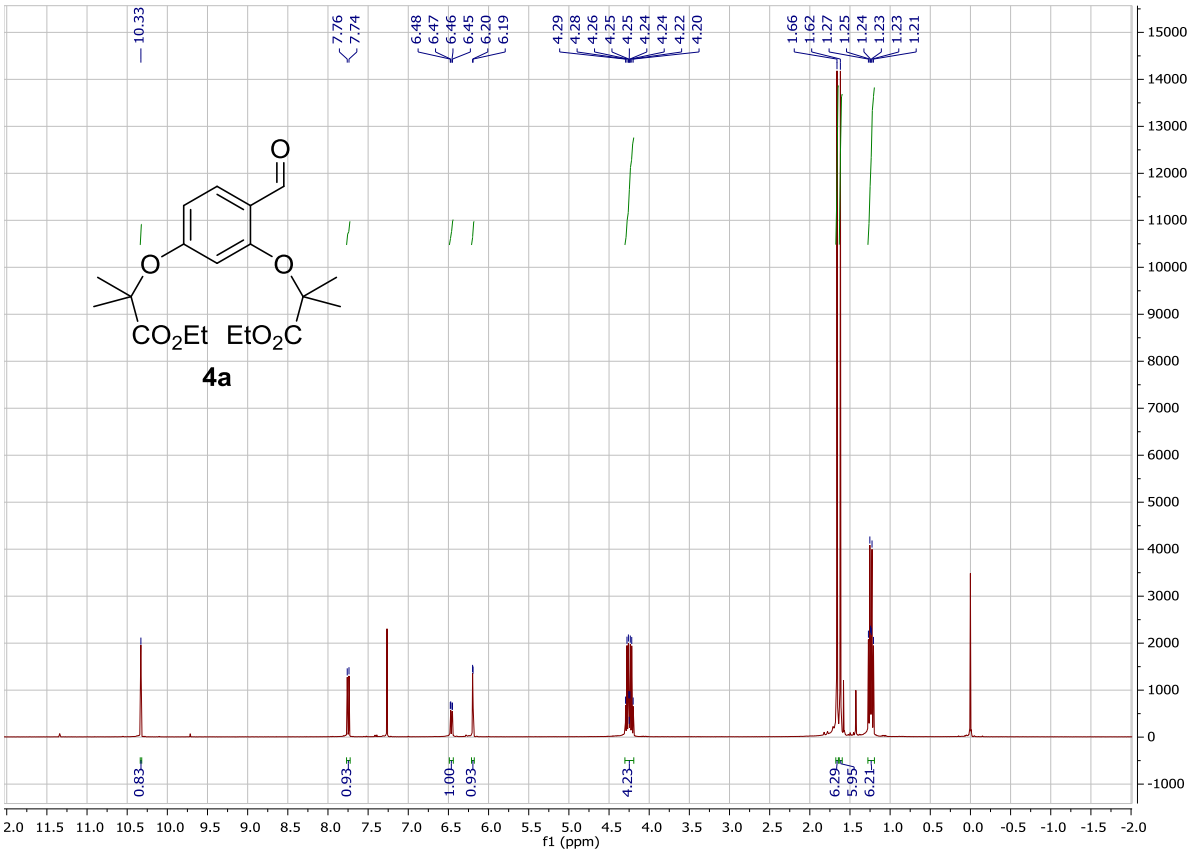
^{1}H NMR (600 MHz, DMSO-d_6)



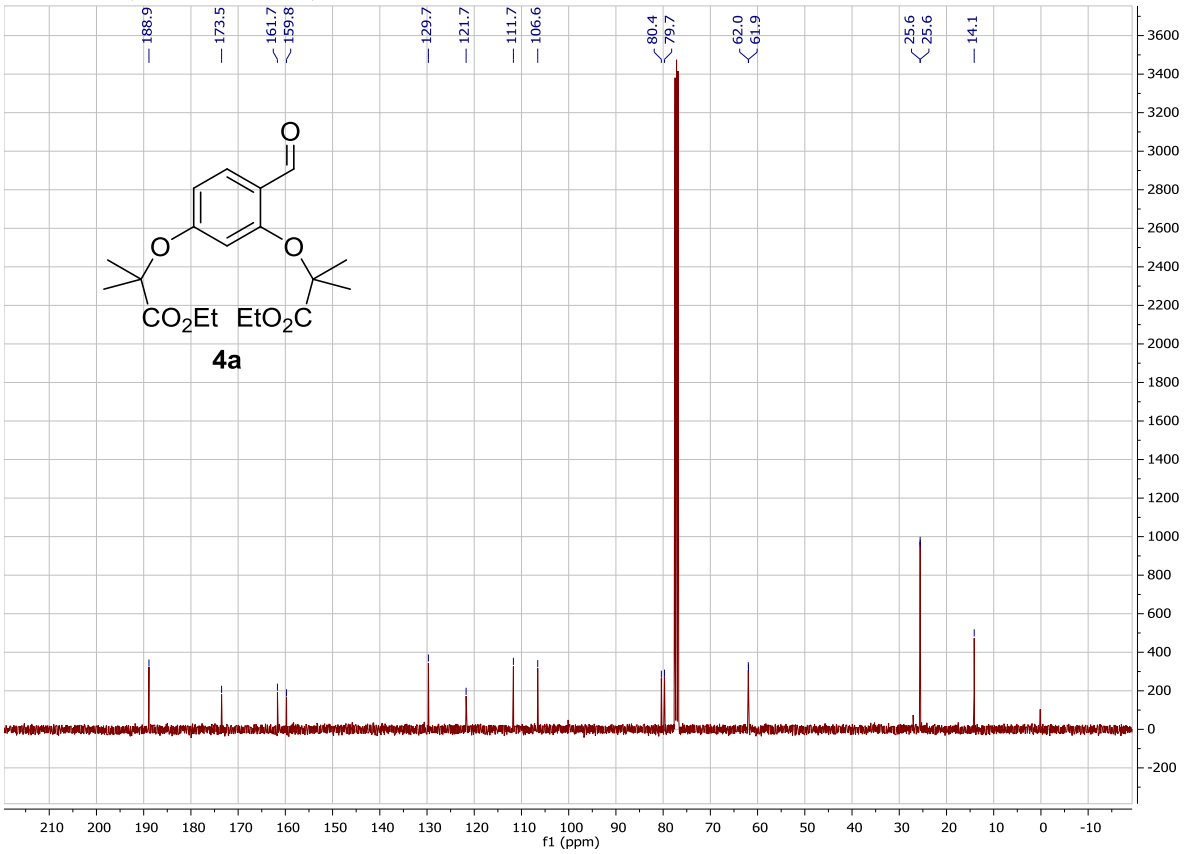
^{13}C NMR (150 Hz, DMSO-d_6)



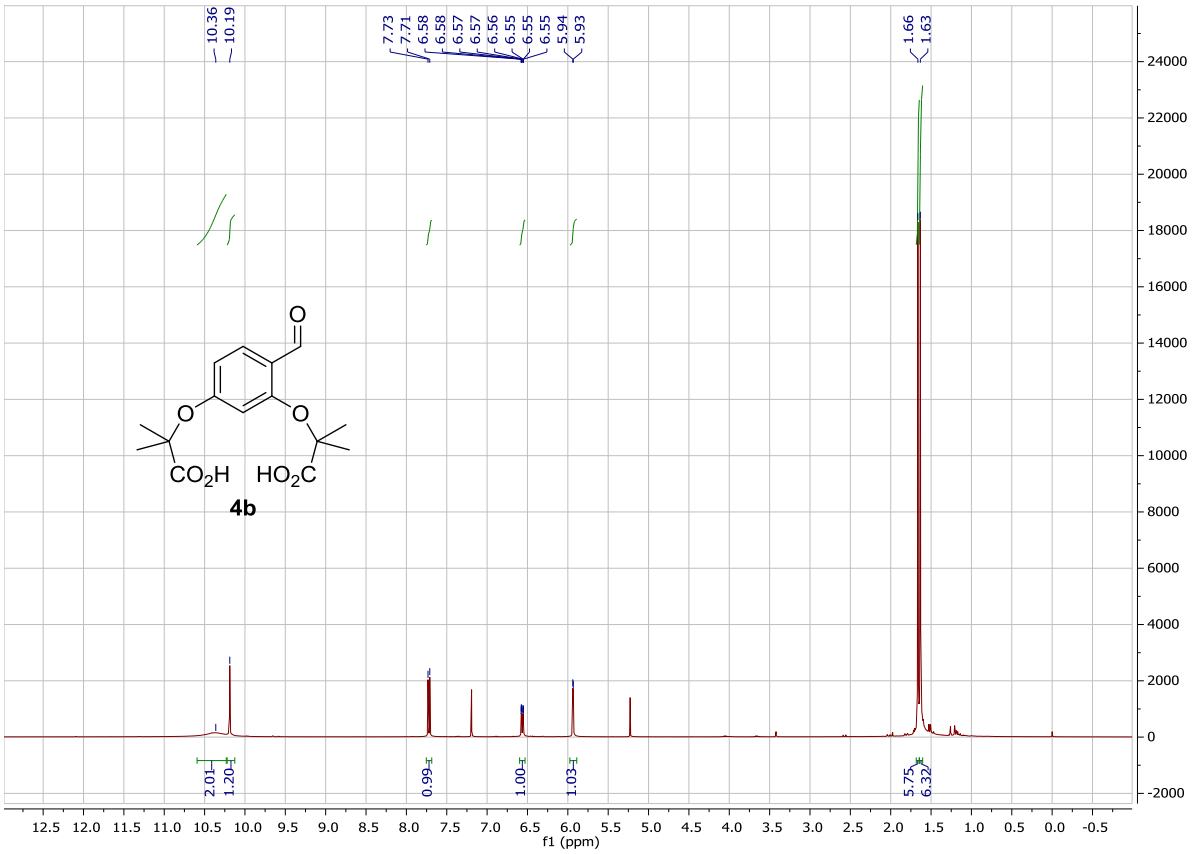
^{1}H NMR (400 MHz, $CDCl_3$)



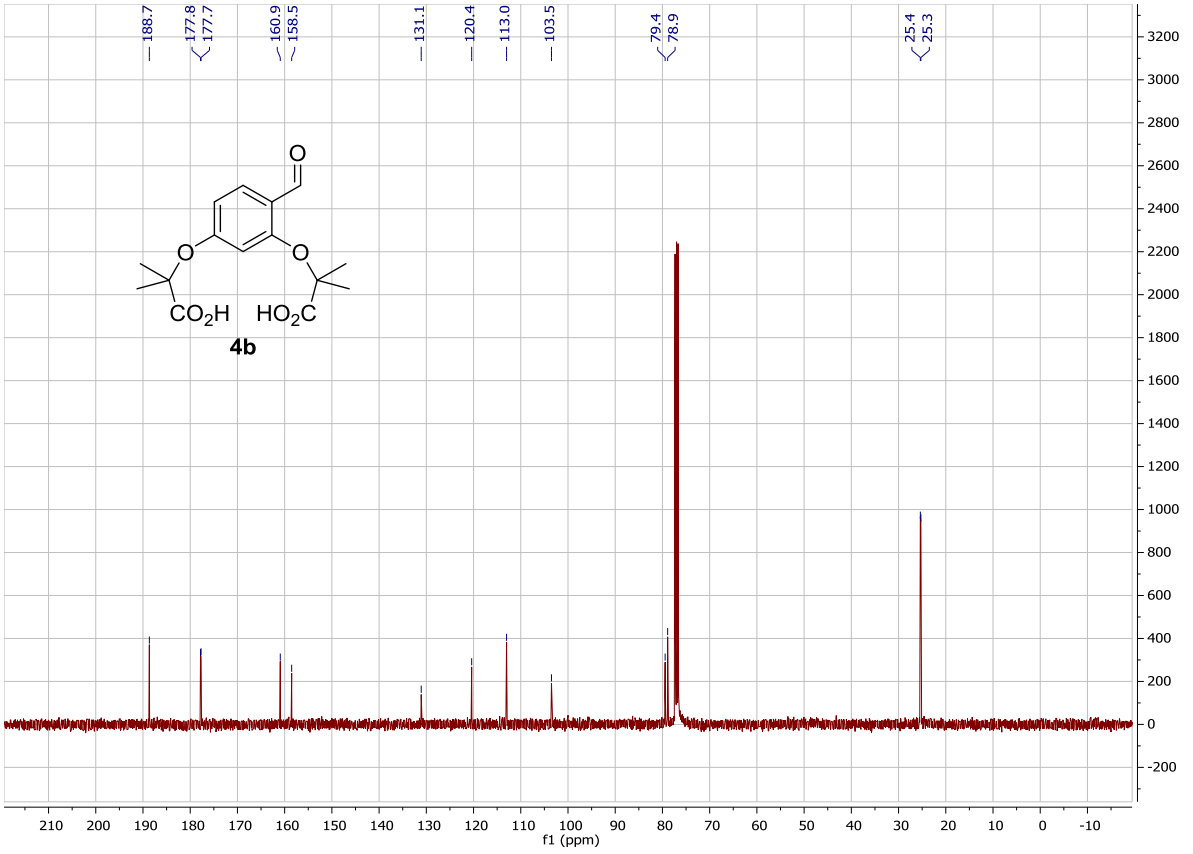
^{13}C NMR (100 Hz, $CDCl_3$)



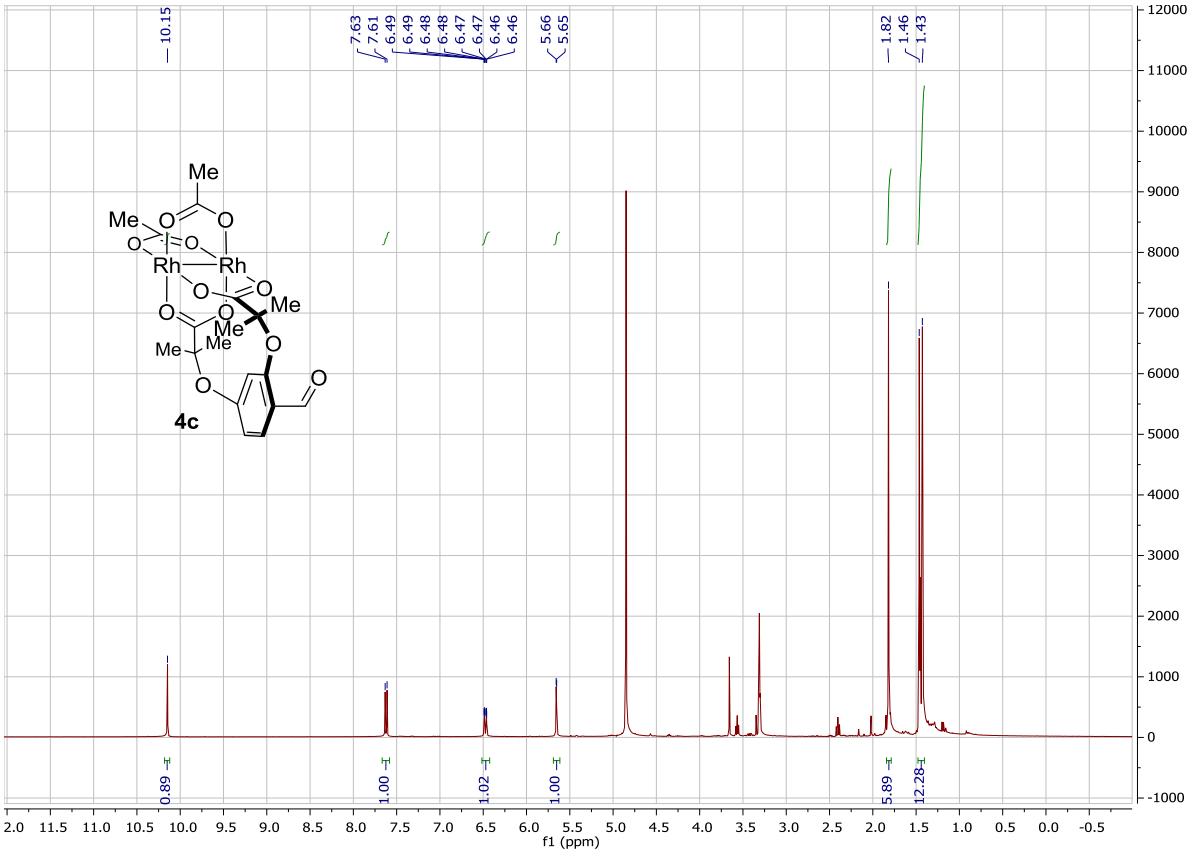
^{1}H NMR (400 MHz, $CDCl_3$)



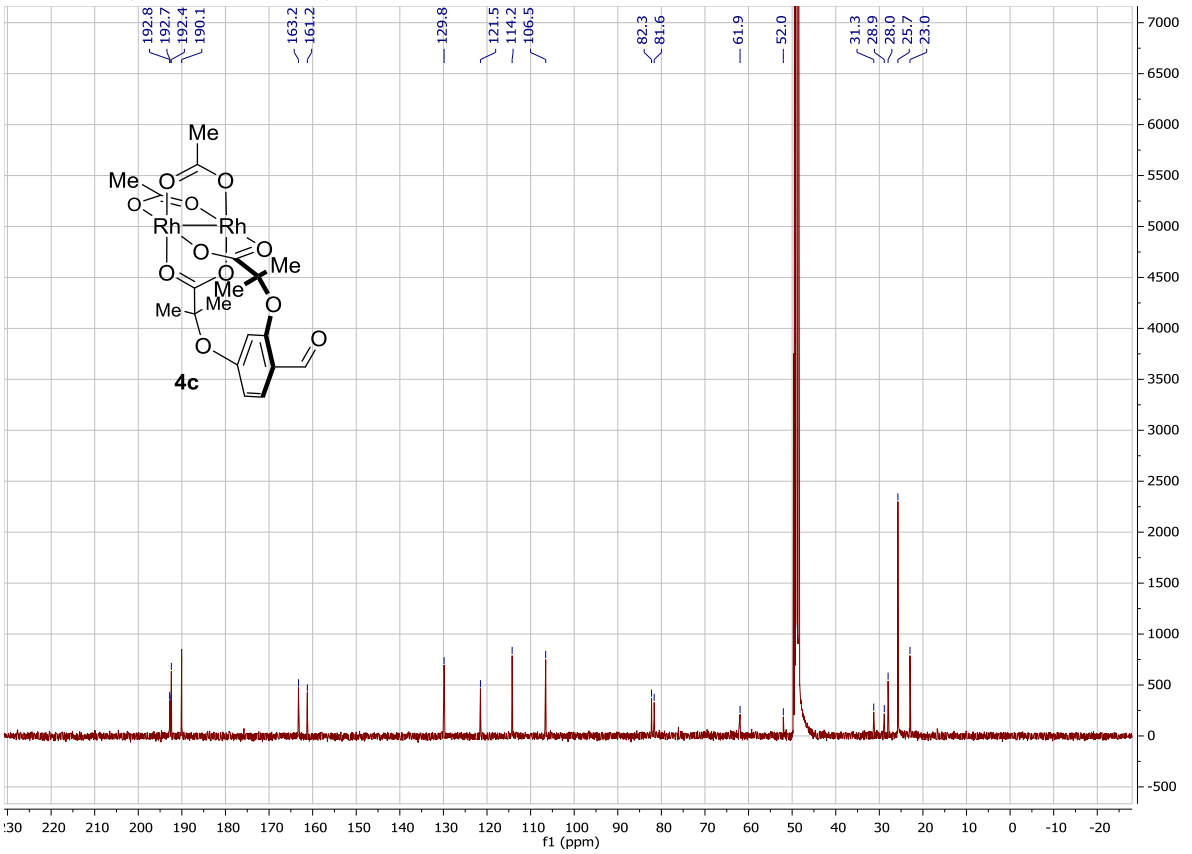
^{13}C NMR (100 Hz, $CDCl_3$)



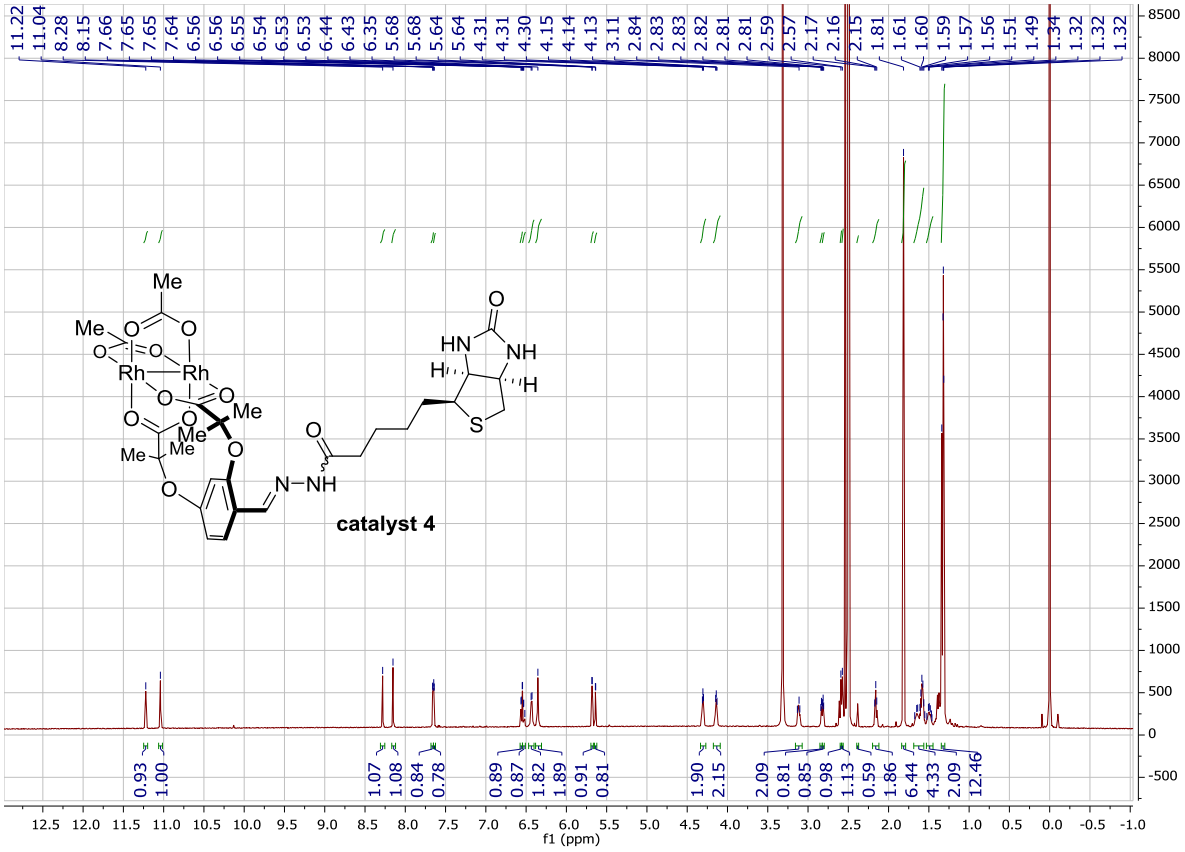
^{1}H NMR (400 MHz, CD_3OD)



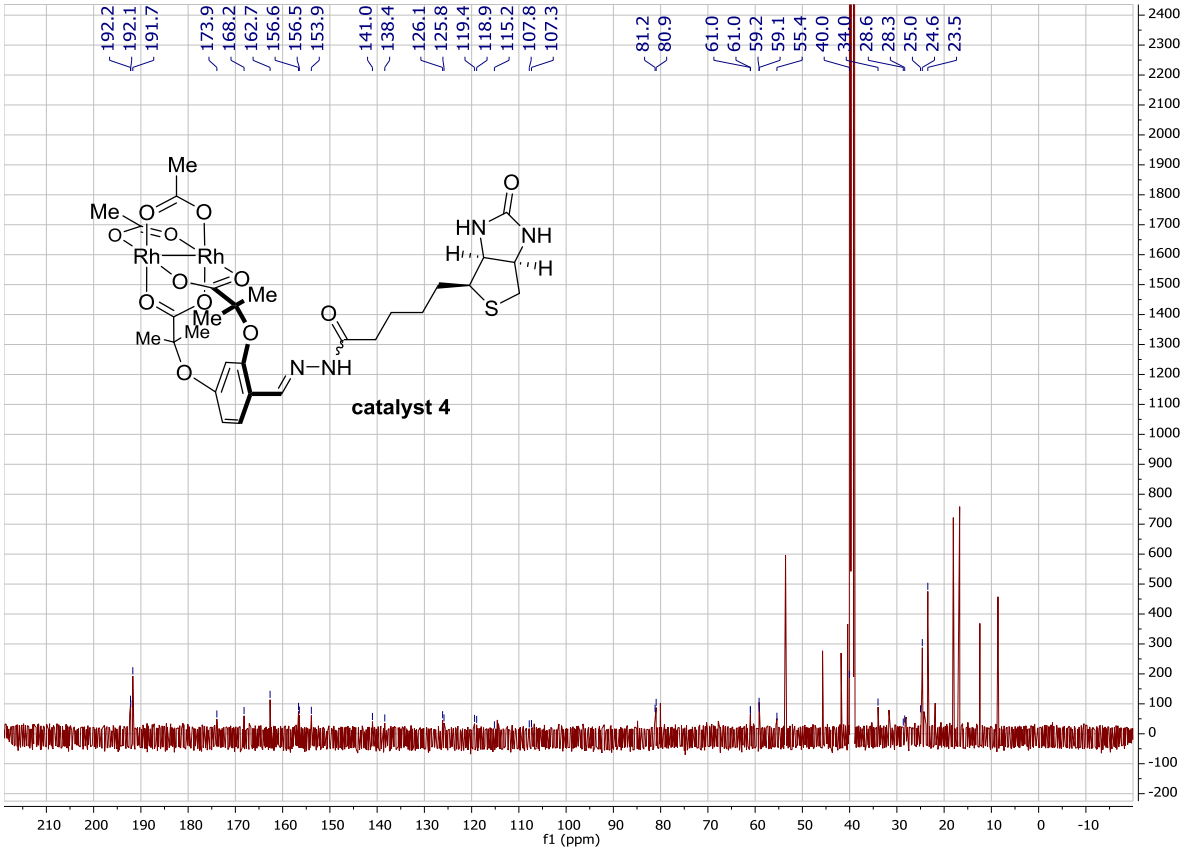
^{13}C NMR (100 Hz, CD_3OD)



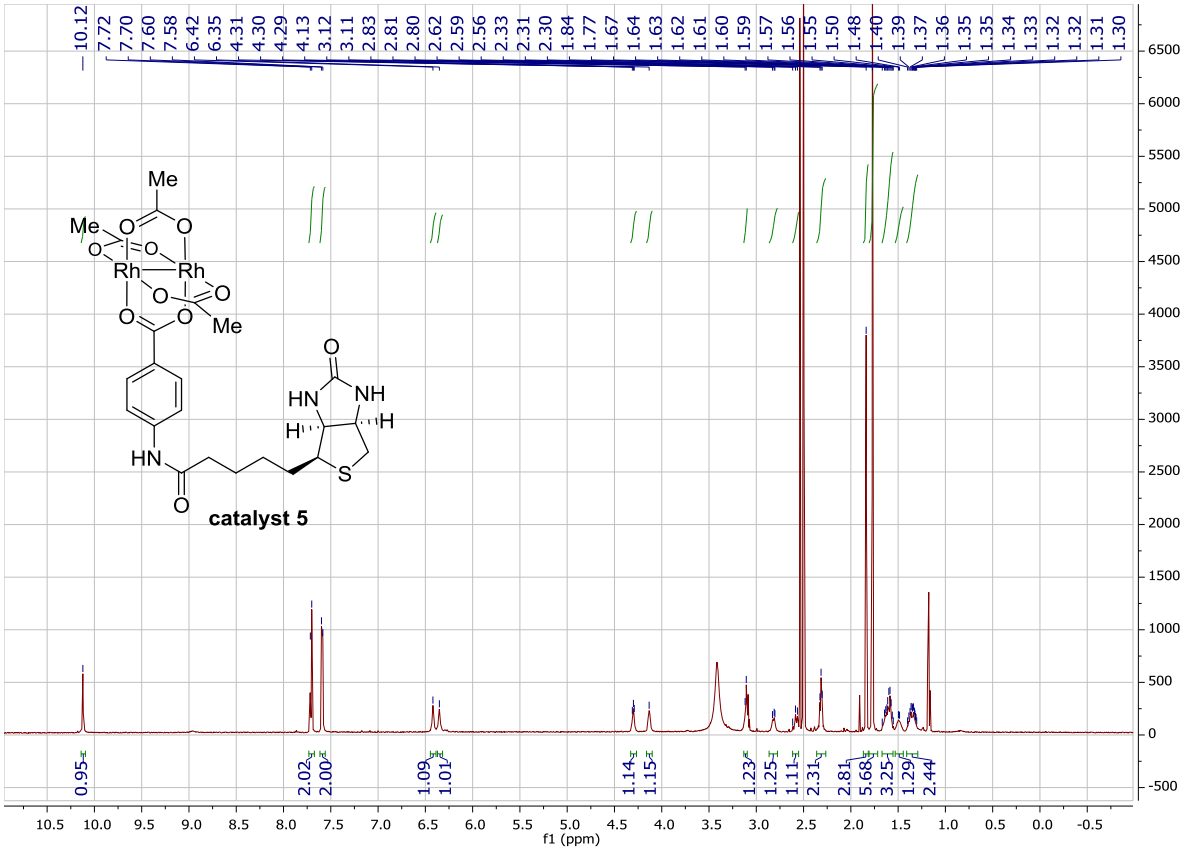
^{1}H NMR (600 MHz, DMSO-d_6)



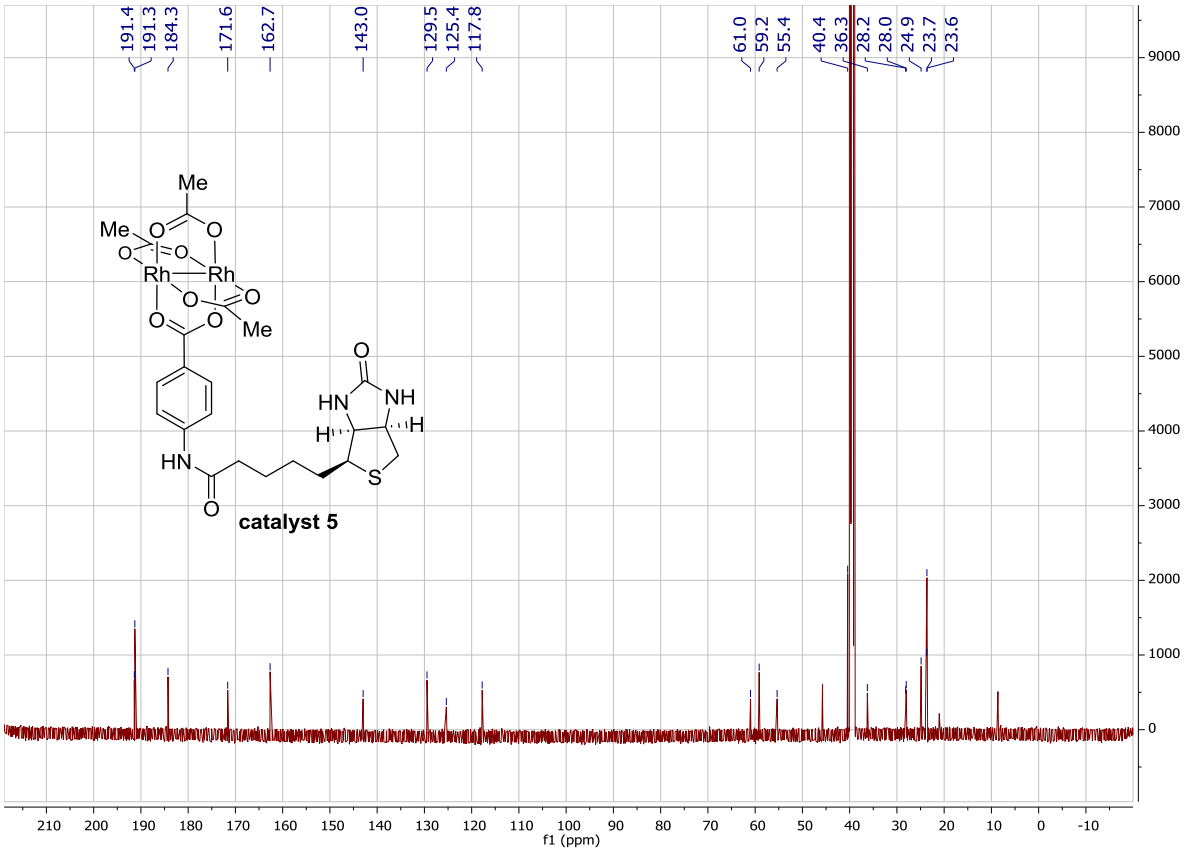
^{13}C NMR (150 Hz, DMSO-d_6)



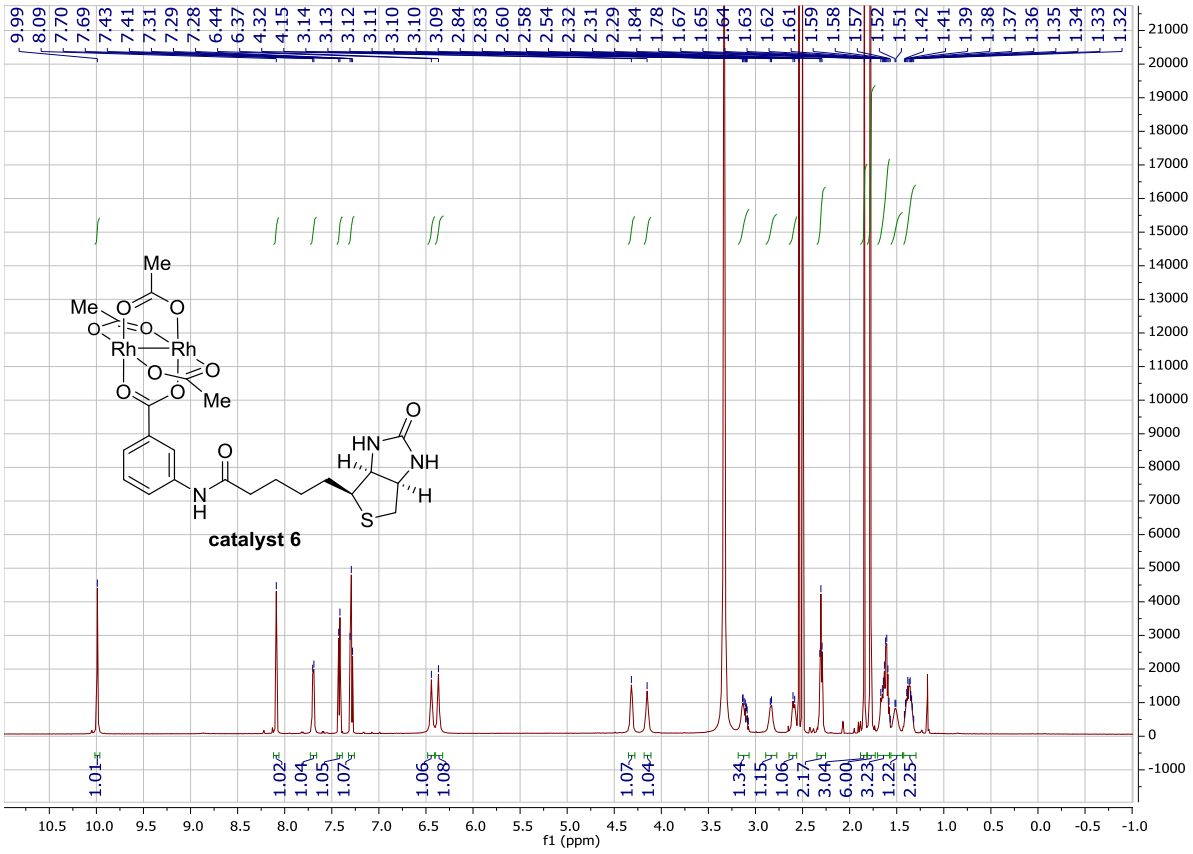
^{1}H NMR (600 MHz, DMSO-d_6)



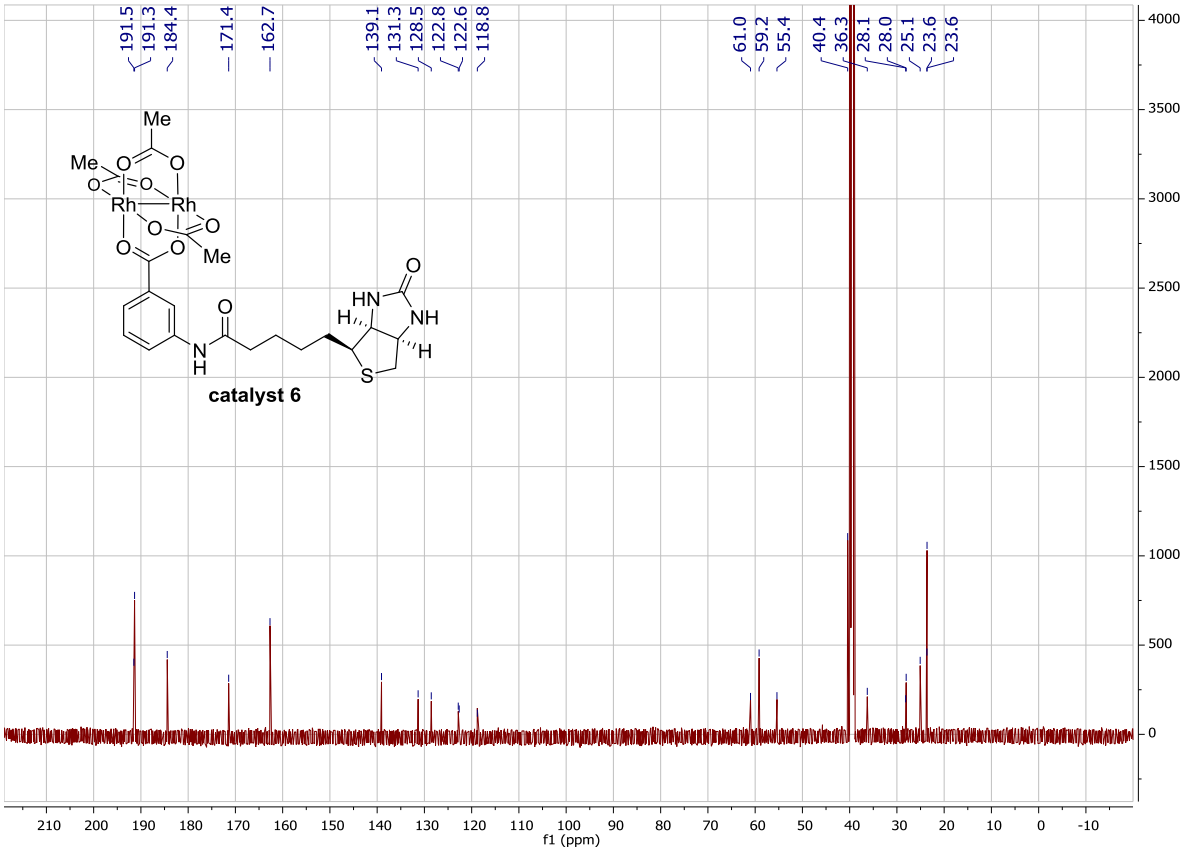
^{13}C NMR (150 Hz, DMSO-d_6)



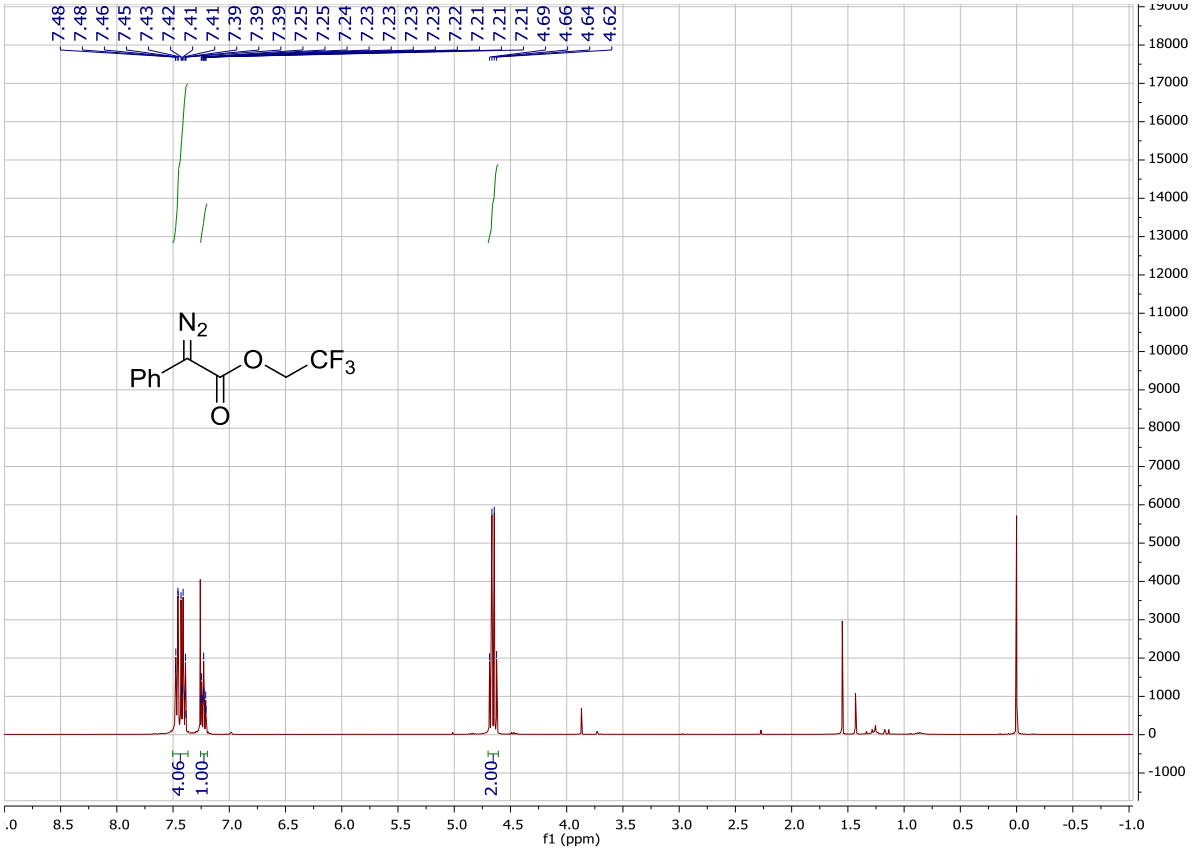
^{1}H NMR (600 MHz, DMSO-d_6)



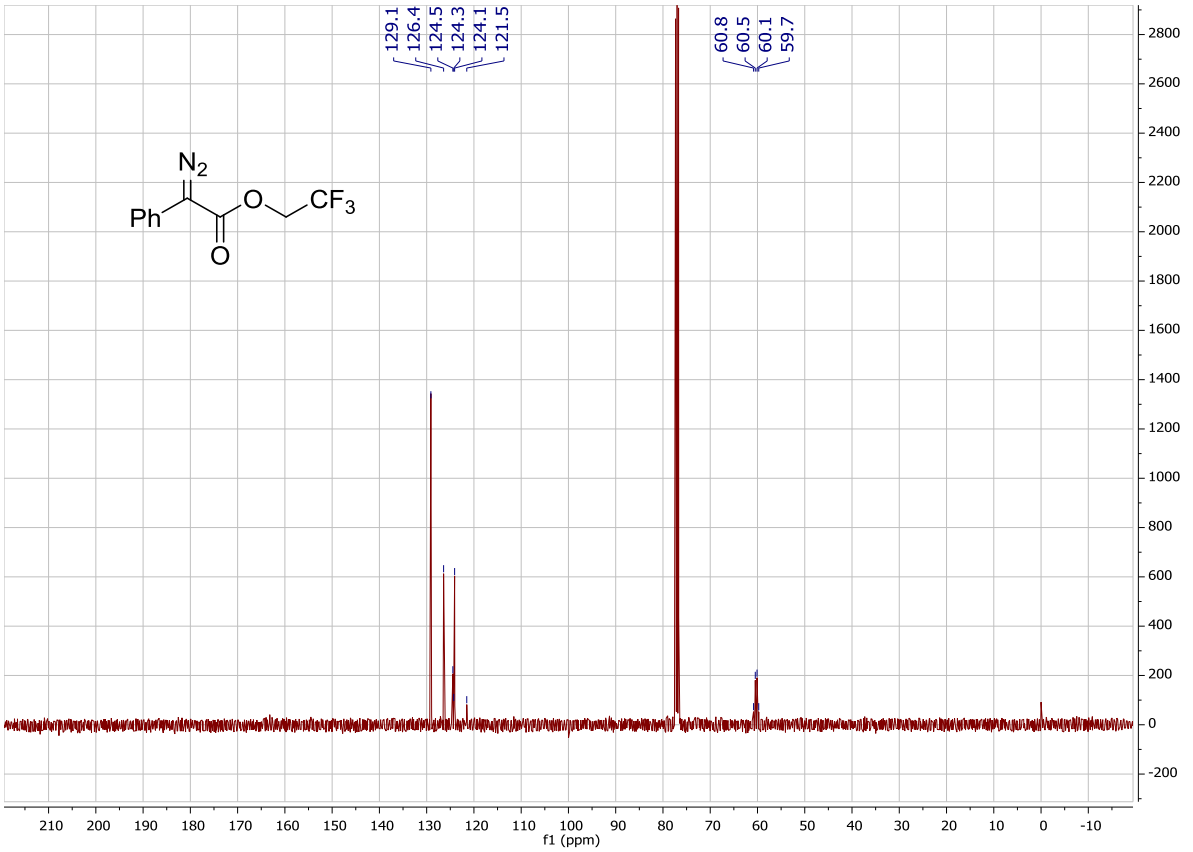
^{13}C NMR (150 Hz, DMSO-d_6)



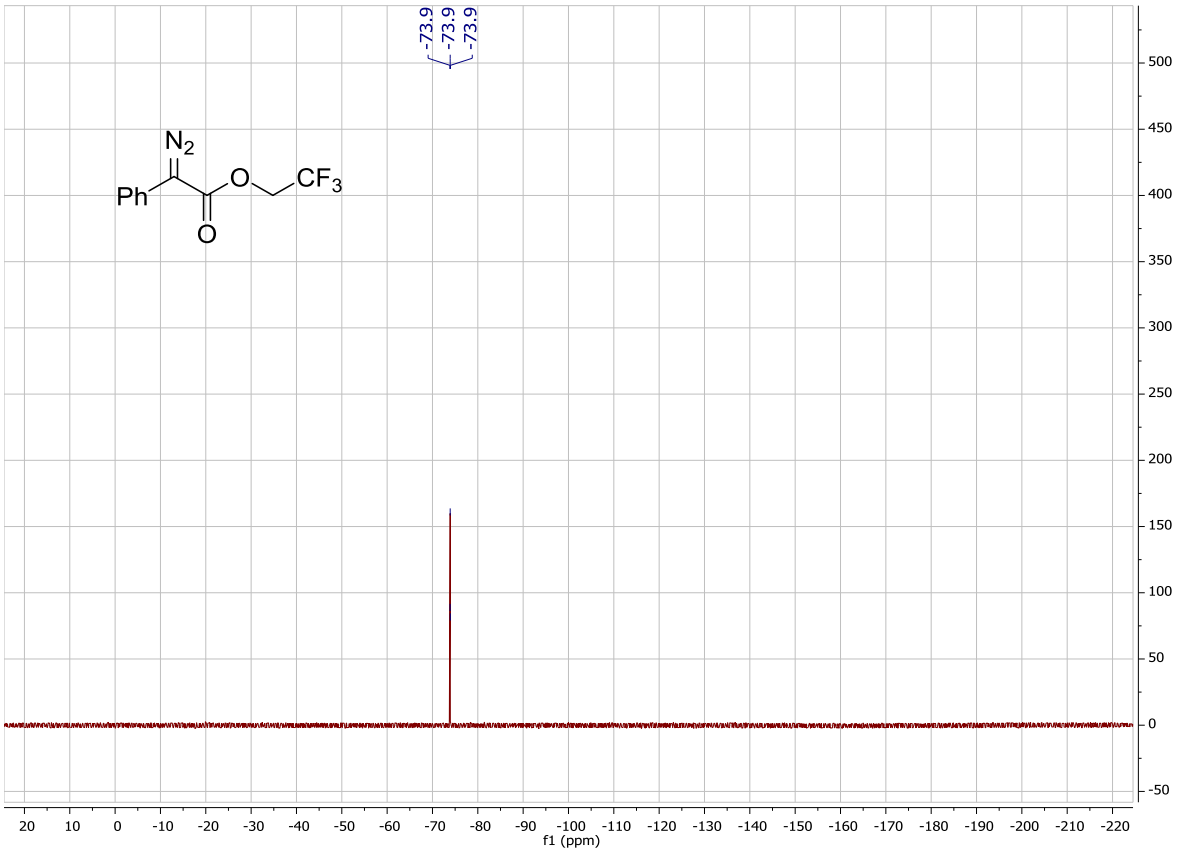
^{1}H NMR (400 MHz, $CDCl_3$)



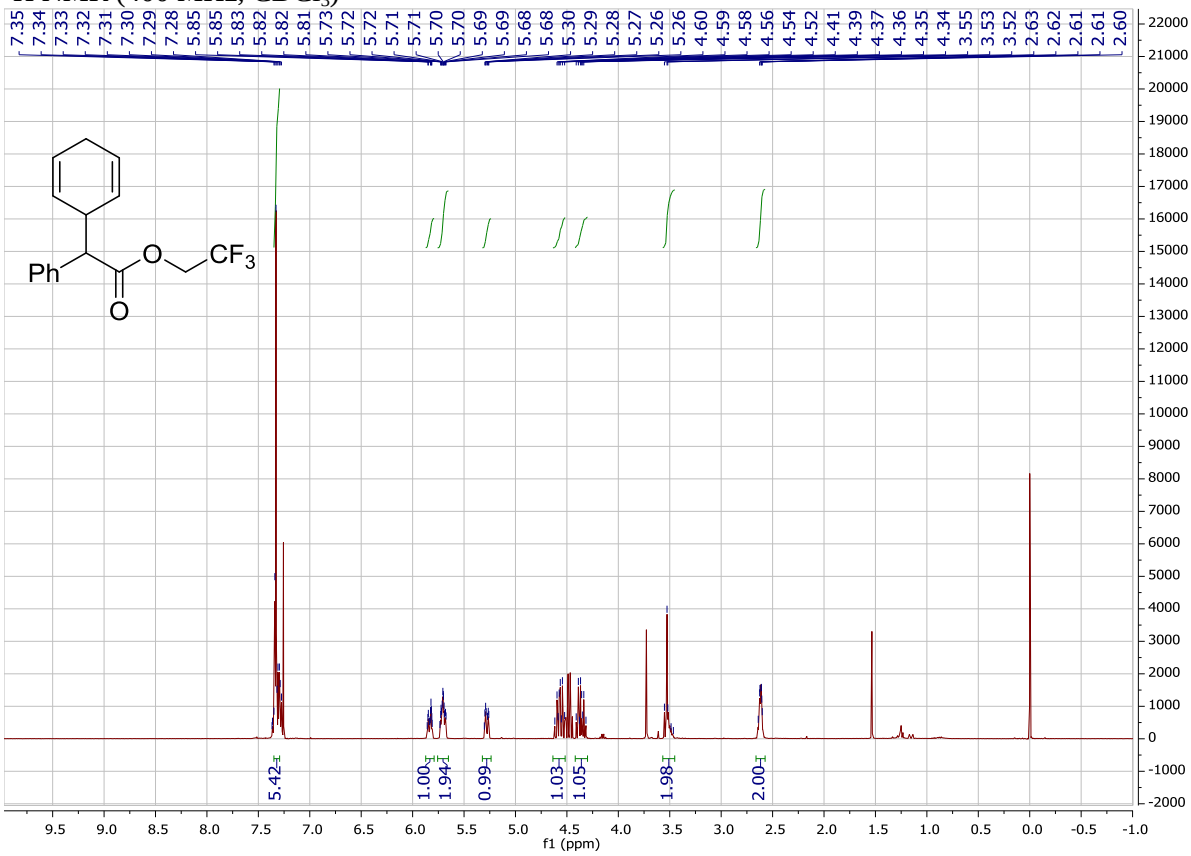
^{13}C NMR (100 Hz, $CDCl_3$)



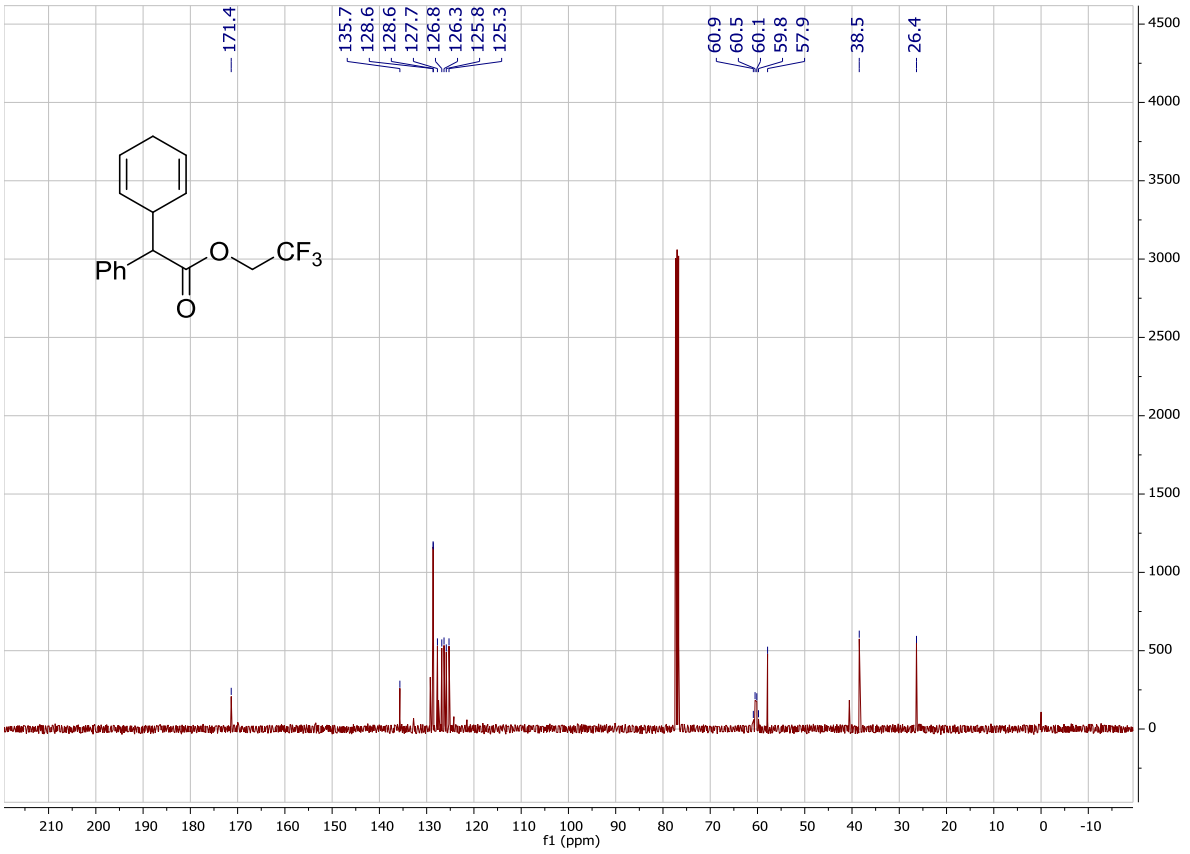
^{19}F NMR (376.5 MHz, $CDCl_3$)



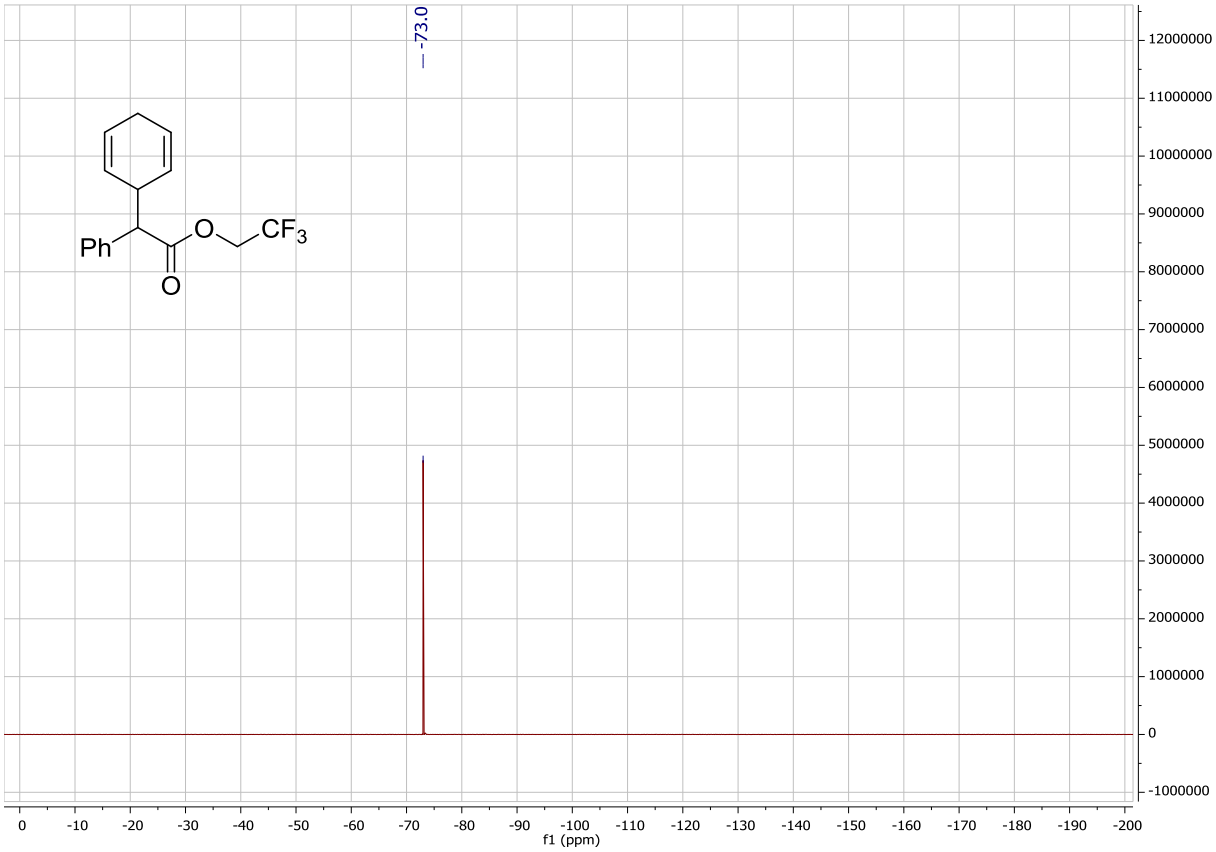
^{1}H NMR (400 MHz, $CDCl_3$)



^{13}C NMR (100 Hz, $CDCl_3$)



^{19}F NMR (376.5 MHz, $CDCl_3$)



3.6 Comments on dirhodium complexes

(a) To form efficient dirhodium-based ArMs, the protein scaffold should have large enough active site to accommodate the large dirhodium mioety. Whether the dirhodium could denature the protein or protein residues could in turn affect the dirhodium need to be carefully evaluated.

(b) The dirhodium moiety should be bound to protein scaffold tightly and at the right distance of the surrounding residues. Directed evolution is a good way to identify hotspots for activity and selectivity improvement. As dirhodium core is stable in cellular environment, using cell lysates or whole cells directly for catalysis may help to increase the thoughput of the screening.

3.7 Reference

(1) C. Qin, V. Boyarskikh, J. H. Hansen, K. I. Hardcastle, D. G. Musaev and H. M. L. Davies, *J. Am. Chem. Soc.*, **2011**, 133, 19198–19204.

(2) C. Werlé, R. Goddard, P. Philipps, C. Farès and A. Fürstner, *Angew. Chemie., Int. Ed.*, **2016**, 55, 10760–10765.

(3) F. G. Adly, M. G. Gardiner and A. Ghanem, *Chem. - Eur. J.*, **2016**, 22, 3447–3461.

(4) H. M. L. Davies and D. Morton, *Chem. Soc. Rev.*, **2011**, 40, 1857.

(5) K. Liao, S. Negretti, D. G. Musaev, J. Bacsa and H. M. L. Davies, *Nature*, **2016**, 533, 230–234.

(6) M. C. Pirrung, H. Liu and A. T. Morehead, *J. Am. Chem. Soc.*, **2002**, 124, 1014–1023.

(7) J.-X. Guo, T. Zhou, B. Xu, S.-F. Zhu and Q.-L. Zhou, *Chem. Sci.*, **2016**, 7, 1104–1108.

(8) B. Xu, S.-F. Zhu, Z.-C. Zhang, Z.-X. Yu, Y. Ma and Q.-L. Zhou, *Chem. Sci.*, **2014**, 5, 1442.

(9) M. B. Minus, M. K. Kang, S. E. Knudsen, W. Liu, M. J. Krueger, M. L. Smith, M. S. Redell and Z. T. Ball, *Chem. Commun.*, **2016**, 52, 11685–11688.

(10) Z. T. Ball, *Curr. Opin. Chem. Biol.*, **2015**, 25, 98–102.

(11) K. Tishinov, K. Schmidt, D. Häussinger and D. G. Gillingham, *Angew. Chemie., Int. Ed.*, **2012**, 51, 12000–12004.

(12) D. G. Bachmann, P. J. Schmidt, S. N. Geigle, A. Chougnet, W. D. Woggon and D. G. Gillingham, *Adv. Synth. Catal.*, **2015**, 357, 2033–2038.

(13) P. Srivastava, H. Yang, K. Ellis-Guardiola and J. C. Lewis, *Nat. Commun.*, **2015**, 6, 1–8.

(14) H. Yang, A. M. Swartz, H. J. Park, P. Srivastava, K. Ellis-Guardiola, D. M. Upp, G. Lee, K. Belsare, Y. Gu, C. Zhang, R. E. Moellering and J. C. Lewis, *Nat. Chem.*, **2018**, 10, 318–324.

(15) M. E. Wilson and G. M. Whitesides, *J. Am. Chem. Soc.*, **1978**, 100, 306–307.

(16) M. Dürrenberger, T. Heinisch, Y. M. Wilson, T. Rossel, E. Nogueira, L. Knörr, A. Mutschler, K. Kersten, M. J. Zimbron, J. Pierron, T. Schirmer and T. R. Ward, *Angew. Chemie., Int. Ed.*, **2011**, 50, 3026–3029.

(17) A. Chatterjee, H. Mallin, J. Klehr, J. Vallapurackal, A. D. Finke, L. Vera, M. Marsh and T. R. Ward, *Chem. Sci.*, **2016**, 7, 673–677.

(18) M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke and T. R. Ward, *Nature*, **2016**, 537, 661–665.

(19) S. I. Mann, T. Heinisch, T. R. Ward and A. S. Borovik, *J. Am. Chem. Soc.*, **2017**, 139, 17289–17292.

(20) A. Pordea, M. Creusa, J. Panek, C. Duboc, D. Mathis, M. Novic and T. R. Ward, *J. Am. Chem. Soc.*, **2008**, 130, 8085–8088.

(21) J. Bos, W. R. Browne, A. J. M. Driessen and G. Roelfes, *J. Am. Chem. Soc.*, **2015**, 137, 9796–9799.

(22) K. Oohora, H. Meichin, Y. Kihira, H. Sugimoto, Y. Shiro and T. Hayashi, *J. Am. Chem. Soc.*,

2017, 139, 18460–18463.

(23) A. G. Jarvis, L. Obrecht, P. J. Deuss, W. Laan, E. K. Gibson, P. P. Wells and P. C. J. Kamer, *Angew. Chemie., Int. Ed.*, **2017**, 56, 13596–13600.

(24) T. Di Meo, W. Ghattas, C. Herrero, C. Velours, P. Minard, J. P. Mahy, R. Ricoux and A. Urvoas, *Chem. - Eur. J.*, **2017**, 23, 10156–10166.

(25) N. Fujieda, T. Nakano, Y. Taniguchi, H. Ichihashi, H. Sugimoto, Y. Morimoto, Y. Nishikawa, G. Kurisu and S. Itoh, *J. Am. Chem. Soc.*, **2017**, 139, 5149–5155.

(26) P. Dydio, H. M. Key, A. Nazarenko, J. Y. E. Rha, V. Seyedkazemi, D. S. Clark and J. F. Hartwig, *Science*, **2016**, 354, 102–106.

(27) P. S. Coelho, E. M. Brustad, A. Kannan and F. H. Arnold, *Science*, **2013**, 307, 307–311.

(28) M. Bordeaux, V. Tyagi and R. Fasan, *Angew. Chemie., Int. Ed.*, **2015**, 54, 1744–1748.

(29) T. Heinisch and T. R. Ward, *Acc. Chem. Res.*, **2016**, 49, 1711–1721.

(30) F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis and T. R. Ward, *Chem. Rev.*, **2018**, 118, 142–231.

(31) Y. Lou, T. P. Remarchuk and E. J. Corey, *J. Am. Chem. Soc.*, **2005**, 127, 14223–14230.

(32) V. M. Robles, M. Dürrenberger, T. Heinisch, A. Lledós, T. Schirmer, T. R. Ward and J. D. Maréchal, *J. Am. Chem. Soc.*, **2014**, 136, 15676–15683.

(33) Z. Chen, B. V. Popp, C. L. Bovet and Z. T. Ball, *ACS Chem. Biol.*, **2011**, 6, 920–925.

(34) M. Jeschek, S. Panke and T. R. Ward, *Methods Enzymol.*, **2016**, 580, 539–556.

(35) M. M. Pellizzoni, F. Schwizer, C. W. Wood, V. Sabatino, Y. Cotelle, S. Matile, D. N. Woolfson and T. R. Ward, *ACS Catal.*, **2018**, 8, 1476–1484.

(36) J. Bickley, R. Bonar-Law, T. McGrath, N. Singh and A. Steiner, *New. J. Chem.*, 2004, **28**, 425-433.

(37) H. Yang, P. Srivastava, C. Zhang, and J. C. Lewis, *ChemBioChem*, 2014, **15**, 223-227.

(38) M. Skander, N. Humbert, J. Collot, J. Gradinaru, G. Klein, A. Loosli, J. Sauser, A. Zocchi, F. Gilardoni and T. R. Ward, *J. Am. Chem. Soc.*, 2004, **126**, 14411-14418.

(39) M. Grünberg and L. J. Gooßen, *Chem. - Eur. J.*, 2013, **19**, 7334-7337.

Chapter 4 | Directed Evolution of Artificial Metalloenzymes for Transfer Hydrogenation of a Self-Immolative Substrate in *E. coli*'s Periplasm

Jingming Zhao, Hendrik Mallin, Christian Trindler, Micehla M. Pellizzoni, Johannes G. Rebelein and Thomas R. Ward*

Department of Chemistry, University of Basel, Mattenstrasse 24a, BPR 1096, Basel, CH-4058, Switzerland.

This work is ready for submission.

4.1 Abstract

Artificial metalloenzymes (ArMs), which combine an abiotic metallocofactor with a protein scaffold, catalyze various synthetically useful transformations. To complement the natural enzymes' repertoire, effective optimization protocols to improve ArM's performance are required. Here we report on our efforts to optimize an artificial transfer hydrogenase's activity (ATHase) using *E. coli* whole cells. For this purpose, we rely on a self-immolative quinolinium substrate surrogate which, upon reduction, releases fluorescent umbelliferone, thus allowing efficient screening. After four rounds of directed evolution in the presence of *E. coli* whole cells, the evolved ATHase displayed up to fivefold increase in its transfer hydrogenation activity compared to the wild-type ATHase.

4.2 Introduction

Directed evolution is a powerful means to optimize the performance of genetically-encoded proteins.[1-5] In recent years, both natural and artificial enzymes have been engineered by directed evolution to catalyze a variety of new-to-nature reactions.[6-14] Artificial metalloenzymes (ArM) result from the introduction of an abiotic metallocofactor within a protein scaffold. In this context, hemoproteins[15,16], streptavidin (Sav hereafter)[7,17] and *Pfu* prolyl-oligopeptidase[11,12] have proven most versatile. The most active ArM reported to date rely on precious metal cofactors, which are frequently poisoned by the presence of thiols.[18-21] Critical analysis of a typical workflow used for the directed evolution of ArMs reveals the following bottlenecks: *i)* need to perform catalysis either using purified protein samples or in a thiol-free environment (i.e. outside the cytoplasm which contains mM concentrations of glutathione in aerobic cells), *ii)* HPLC analysis of the reaction outcome is often time-consuming.
Herein, we report on our efforts to address both the above challenges by compartmentalizing an artificial transfer-hydrogenase (ATHase hereafter) in the periplasm of *E. coli* and relying on a self-immolative substrate surrogate that releases a fluorophore upon reduction of its iminium moiety.

4.3 Results and discussion

Self-immolative substrates consist of a cap, a self-immolative core and a leaving group.[22] A catalytic reaction modifies the cap which, in turn, triggers the release of the leaving group, via the intermediary of the self-immolative core, Figure 1a.[23] Self-immolative substrates have found wide applications in bioanalysis[24], prodrugs[25,26], high-throughput screening or biomolecular imaging[27] etc. We hypothesized that we could design a quinoline substrate that, upon N=C reduction, would undergo self-immolation to release a fluorophore via the formation of iminoquinone-methide intermediate,[28] Figure 1b. Although several anchoring strategies and protein scaffolds have been exploited for the

creation of ATHase,[29-31] we selected the biotin-streptavidin couple to ensure localization of the Ir-based cofactor. Analysis of the various X-ray structures [Cp*Ir(biot-*p*-L)Cl] · Sav,[32,33] reveals that the cofactor, located in the biotin-binding vestibule, is solvent exposed. We reasoned that introduction of additional structural elements around the biotin-binding vestibule may have a significant impact on the catalytic performance.[34] Thanks to its stability and plasticity, we have shown that Sav is remarkably tolerant towards the modification of its loops.[35] Having repeatedly reported that mutations within the 7,8-loop (residues between 112 and 124) had the strongest effect on catalytic performance of various ArM based on the biotin-streptavidin technology,[36,37] we selected a 24-residue helix-turn-helix motif (foldit player design, FPD) designed by Baker group[38] (Figure 2a), and set out to insert it between residues 115 and 117 of Sav WT, Figure 2. The structure of the [(Cp*Ir(biot-*p*-L)Cl] · Sav-FPD was modelled using Rosetta.[39] To our delight, the FPD-chimera could be expressed as a soluble fraction in *E. coli* and purified by affinity chromatography on an iminobiotin-sepharose matrix, thus highlighting its biotin-binding affinity (See SI for details). To circumvent the irreversible poisoning of the precious metal cofactor by glutathione,[40,41] we set out to secrete the Sav-FPD chimera to the periplasm by fusing it with the OmpA *N*-terminal signal peptide. SDS-PAGE analysis of *E. coli* cells containing the OmpA-Sav-FPD construct confirmed the localization of the soluble biotin-binding protein within its periplasm (Figure SI_Figure 1). As *E. coli*'s periplasm contains significantly lower glutathione concentrations than in its cytoplasm, and the equilibrium lies mostly to the oxidized disulfide form, we anticipated that the assembled [(Cp*Ir(biot-*p*-L)Cl] · Sav-FPD would maintain its ATHase activity in the periplasm, as we have recently demonstrated for metathesis using a biotinylated Ru-cofactor.[40]

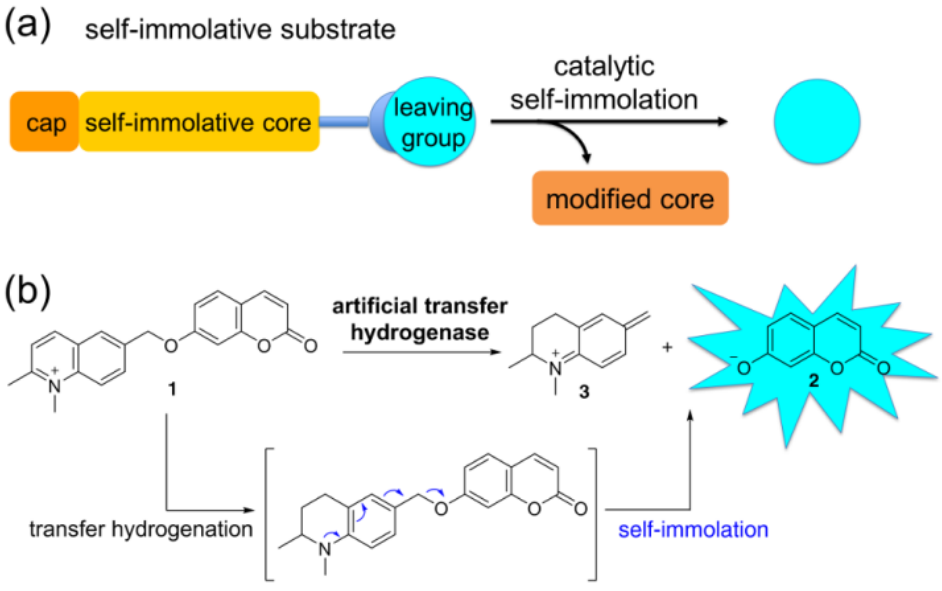


Figure 1. (a) General concept of a self-immolative substrate. (b) Reduction of the quinolinum moiety in substrate **1** leads to the release of umbelliferone **2** which can be detected by fluorescence.

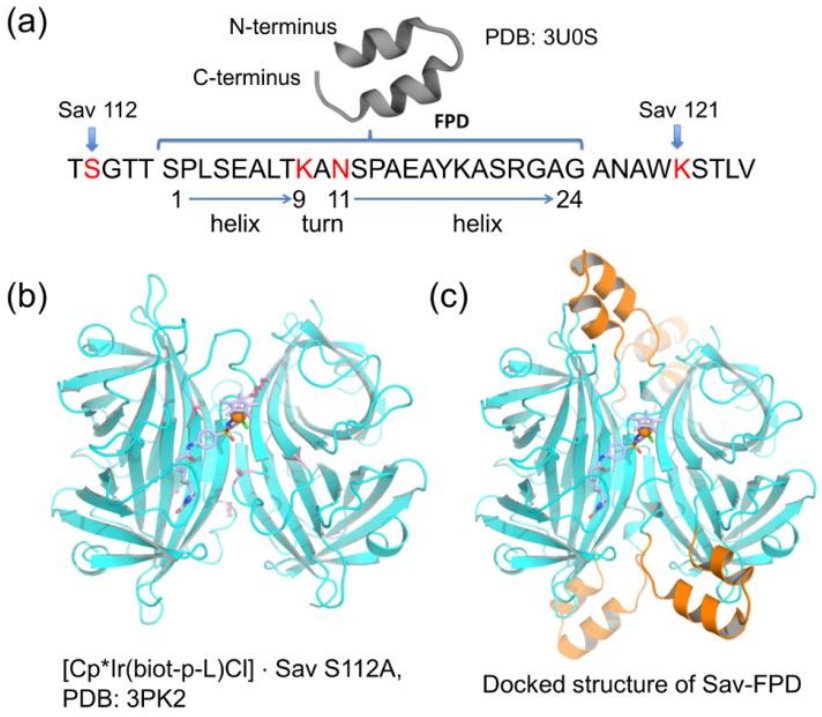


Figure 2. Introducing an additional structural motif around the biotin-binding vestibule may influence ATHase activity. (a) Cartoon representation of the FPD loop and its sequence; the residues highlighted in red were subjected to saturation mutagenesis. (b) Crystal structure of Sav S112A assembled with [Cp*Ir(biot-*p*-L)Cl] · Sav complex (PDB: 3PK2). (c) Modelled structure [Cp*Ir(biot-*p*-L)Cl] · Sav-FPD.

Following the *E. coli* culture using a modified ZYM5052 rich medium, the cells were incubated (30 min on the ice) with the buffer containing [(Cp*Ir(biot-*p*-L)Cl] and washed twice to remove the unbound cofactor (See SI for details). Next, the cell pellet was resuspended in a solution containing formate (1 M) and the substrate **1** (1 mM). Unfortunately, no conversion could be detected after 16 hours at 25 °C, either by fluorescence (322 nm excitation, 440 nm fluorescence of umbelliferone **2**) or by UPLC-MS. Speculating that the low concentrations of glutathione present in the periplasm may suffice to irreversibly poison [(Cp*Ir(biot-*p*-L)Cl], we added diamide (2 mM) to fully oxidize glutathione. Again here, no conversion could be observed. Inspired by the Cu(II)-catalyzed glutathione oxidation,[42,43] we treated the *E. coli* cells with [Cu(gly)$_2$] (2 mM), both at harvest and in the catalysis buffer. To our delight, the uncaging of umbelliferone **2** was observed when this protocol was applied. Thanks to the good Sav-FPD expression levels in *E. coli*'s periplasm, and the sensitivity of the fluorescence-based monitoring of catalysis, the entire screening procedure could be carried out in a 96-well plate format, Figure 3.

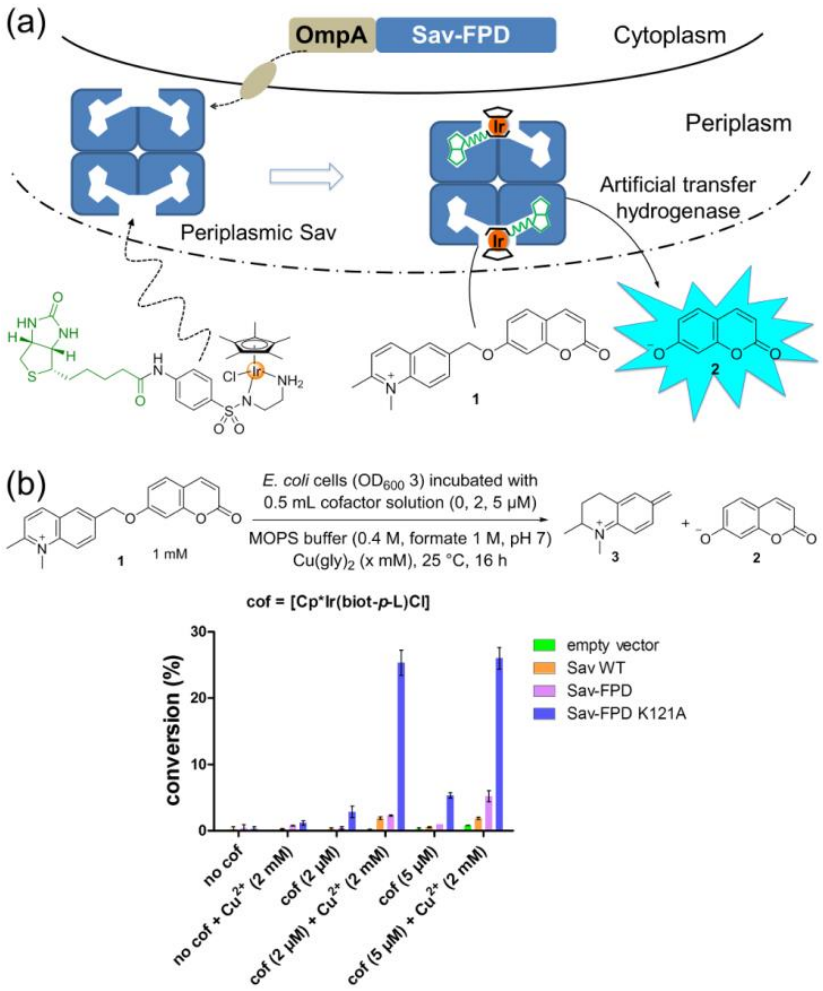


Figure 3. Assembling and screening ATHase in *E. coli* whole cells. (a) Localization of Sav-FPD chimeras within the periplasm of *E. coli* allows to screen whole cells for ATHase activity upon incubation with [Cp*Ir(biot-*p*-L)Cl] using self-immolative quinolinium substrate surrogate **1**. (b) Addition of [$Cu(gly)_2$] significantly improves ATHase activity in *E. coli*.

Past experience with [(Cp*Ir(biot-*p*-L)Cl] · Sav highlighted the detrimental influence of the K121 residue on ATHase activity.[21,44] We hypothesized that this residue may also affect catalytic activity of [(Cp*Ir(biot-*p*-L)Cl] · Sav-FPD. Gratifyingly, upon addition of 2 µM cofactor and 2 mM [$Cu(gly)_2$], the catalytic performance of [(Cp*Ir(biot-*p*-L)Cl] · Sav-FPD K121A was markedly improved compared to [(Cp*Ir(biot-*p*-L)Cl] · Sav-FPD (5% conversion and 26% conversion respectively after 16 hours), Figure 3. This result prompted us to optimize the catalytic activity by directed evolution. We thus built a mutant library relying on the use NHT codons (encode amino acids A, N, D, H, I, L, F, P, S, T, Y, V) for mutations at the following positions: S112, K121, FPD(K9) and FPD(N11), Figure 2a. The NHT primers code mostly for hydrophobic residues, which have most often been shown to positively influence the catalytic performance of ArM based on the biotin-streptavidin technology.[36,37]

In order to ensure full coverage at positions S112 and K121 (i.e. 12 · 12 = 144 mutants), we mutated simultaneously both positions. Accordingly, 450 clones were screened and the best hits were

sequenced. Eight replicates were tested independently revealing that Sav-FPD S112V-K121A performs best. It was thus selected as the template for the next generation for FPD-centered mutagenesis. The positions K9 and N11, which are part of the FPD loop, were randomized sequentially using NHT codons. Unfortunately, none of the triple or quadruple mutants screened outperformed [(Cp*Ir(biot-*p*-L)Cl] · Sav-FPD S112V-K121A chimera for the reduction of the quinolinium substrate surrogate **1** after screening 600 clones, Figure 4.

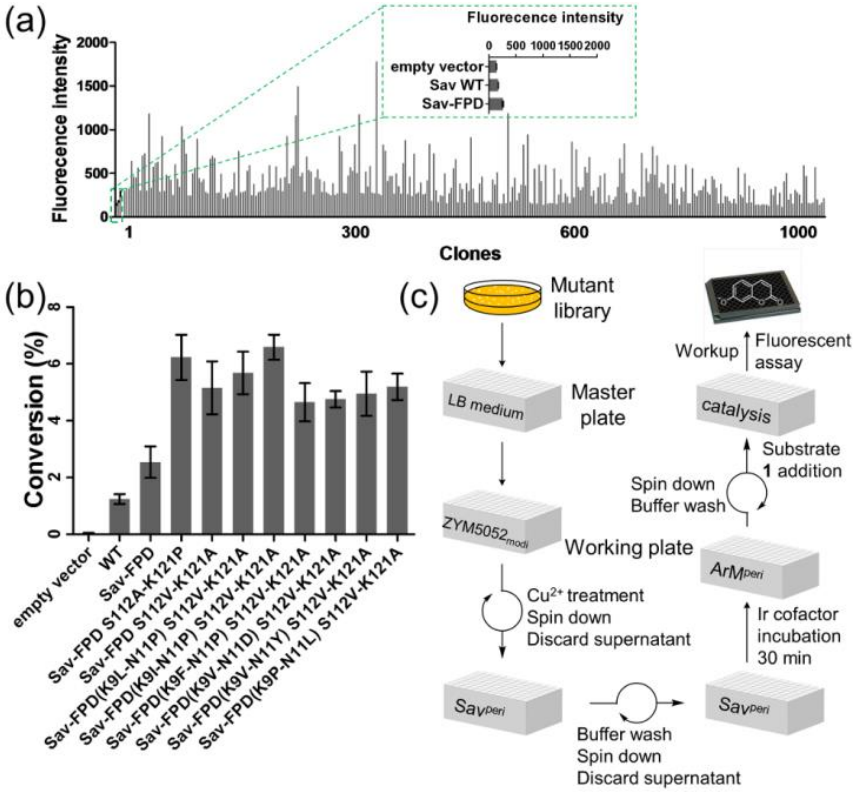


Figure 4. Directed evolution of ATHase using quinolinium substrate **1**. (a) Summary of the streamlined screening protocol for random mutagenesis at positions S112X and K121Y using NHT codons. (b) 8-replicate comparison for the most promising mutants (c) 96-well plate directed evolution protocol.

The observed whole cell ATHase improved activity may result from a combination of two factors: i) improved activity of the Sav-FPD chimera and/or ii) higher protein production levels of the recombinant protein in the periplasm. Screening with purified Sav samples in vitro, allows to deconvolute both these factors and identify the Sav-FPD's activity improvement resulting from the directed evolution. Accordingly, the in vitro ATHase activity towards four imine substrates **4a**-**7a** were tested and the results are summarized in Figure 5.

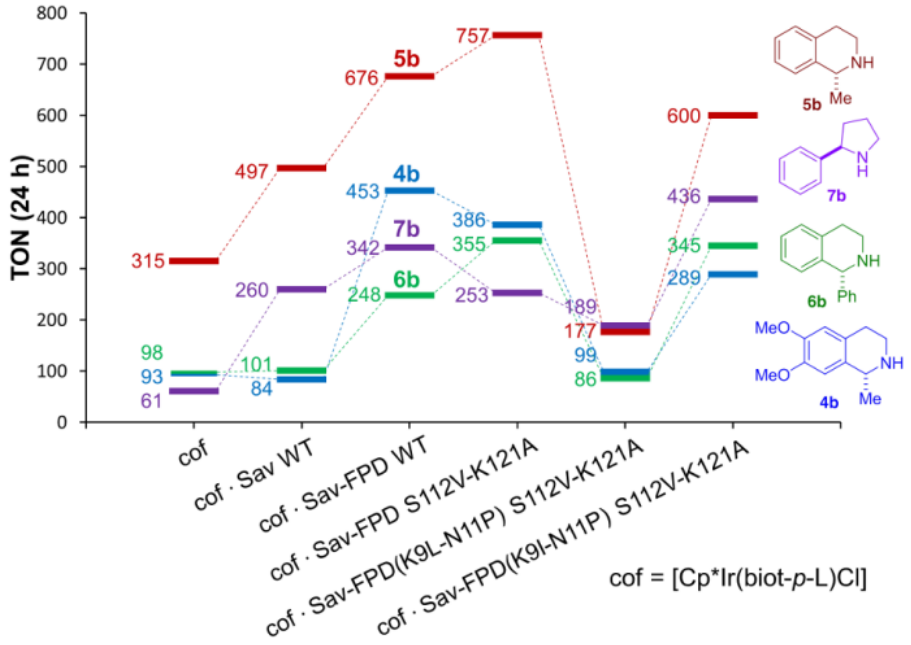


Figure 5. Substrate scope for ATHase using mutants identified using self-immolative substrate **1**. Conversions and ee where determined either by HPLC (substrates **4a**, **5a**, **6a**) or GC (for **7a**) (See SI for details).

The results of the screening using purified Sav-FPD samples with prochiral substrates **4a**-**7a** reveal several striking features: i) Compared to Sav WT, introduction of the FPD structural motif (Sav-FPD) has a positive impact on the turnover numbers after 24 hours (TON). Up to a fivefold higher TON is observed for the salsolidine precursor **4a**. ii) Introduction of mutations at the base of loop 7,8 (i.e. Sav-FPD S112V-K121A) positively influences TON for three imine substrates: **5a**, **6a** and **7a**. iii) These mutations also significantly influence the enantioselectivity of the reduction, especially in the case of imine **4a**. It should be emphasized that the screening protocol focused on improving TON, not ee. iii) Mutations within the FPD motif have ambivalent effects: while Sav-FPD(K9L-N11P) S112V-K121A leads to a significant decrease in TON for all substrates, the Sav-FPD(K9I-N11P) S112V-K121A affords good TON in most cases.

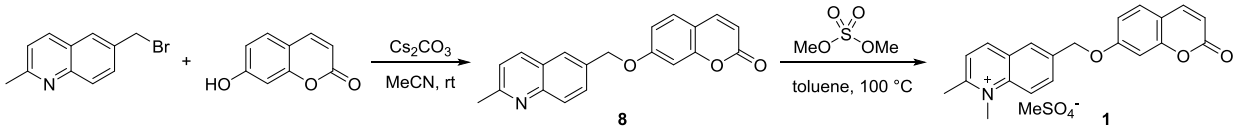
4.4 Conclusion

With the aim of developing a streamlined protocol for the directed evolution of ATHases using *E. coli* whole cells, we report herein three important developments. i) Asymmetric transfer hydrogenation using the [(Cp*Ir(biot-*p*-L)Cl] · Sav ArM can be performed in the periplasm of *E. coli*, upon addition of [Cu(gly)$_2$]. We hypothesize that this contributes to oxidize the glutathione present in the periplasm. ii) The introduction of a 24 aminoacid helix-loop-helix in the immediate proximity of the biotin-binding vestibule has a positive effect of the catalytic performance of [(Cp*Ir(biot-*p*-L)Cl] · Sav-FPD as reflected by the increased TON (up to fivefold improvement). iii) The performance of the ATHase can be improved by directed using *E. coli* whole cells. Upon relying on the self-immolative

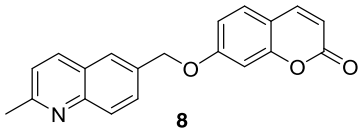
quinolinium substrate **1**, the entire screening protocol can be carried out in a 96 well plate format, from protein expression to fluorescent readout. Current efforts are aimed at characterizing by X-ray the evolved [(Cp*Ir(biot-*p*-L)Cl] · Sav-FPD and adapting this protocol to improve both the activity and the (pseudo)-enantioselectivity of ATHase based on the biotin-streptavidin technology.

4.5 Supporting information

Synthesis of quinolinium substrate 1



7-((2-methylquinolin-6-yl)methoxy)-2H-chromen-2-one 8



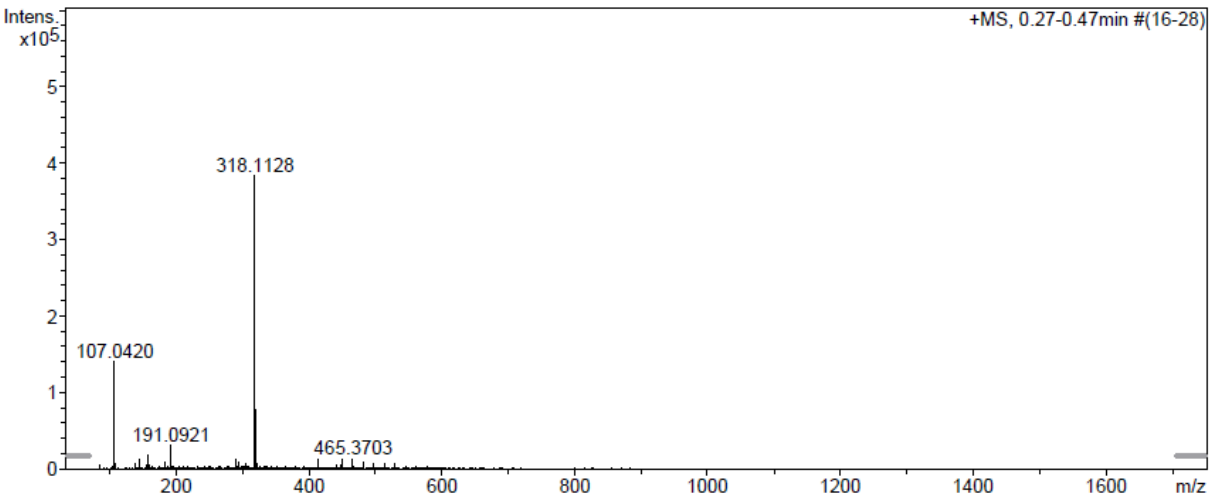
8

To a 50 mL two-necked round-bottom flask, umbelliferone (107 mg, 0.66 mmol, 1.1 equiv.) and Cs_2CO_3 (342 mg, 1.05 mmol, 1.75 equiv.) were added. The flask was evacuated and charged with N_2 three times. Under a mild N_2-flow, MeCN (30 mL) was added and the reaction mixture was stirred at rt for 20 min. Several small portions of 6-(bromomethyl)-2-methylquinoline (142 mg, 0.6 mmol, 1 equiv.) were slowly added into the flask and the mixture was stirred for another 5 h. The solvent was removed under reduced pressure, and the resulting crude mixture was purified by flash chromatography (EtOAc / cyclohexane = 1 / 1 ~ 2 / 1) to afford product **8** (137.5 mg, 72% yield) as white solid.

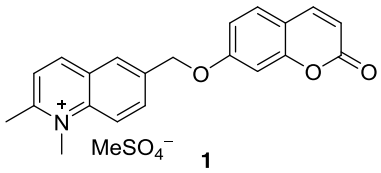
^{1}H NMR (400 MHz, DMSO-d_6) δ 8.28 (d, J = 8.4 Hz, 1H), 8.03 – 7.95 (m, 3H), 7.80 (dd, J = 8.7, 2.0 Hz, 1H), 7.66 (d, J = 8.6 Hz, 1H), 7.45 (d, J = 8.4 Hz, 1H), 7.15 (d, J = 2.4 Hz, 1H), 7.08 (dd, J = 8.6, 2.4 Hz, 1H), 6.30 (d, J = 9.5 Hz, 1H), 5.41 (s, 2H), 2.67 (s, 3H).

^{13}C NMR (100 Hz, DMSO-d_6) δ 161.4, 160.2, 159.0, 155.3, 146.8, 144.3, 136.2, 133.6, 129.5, 129.2, 128.4, 126.6, 125.9, 122.5, 113.0, 112.6, 112.6, 101.7, 69.6, 24.8.

HRMS (ESI, pos.) *m/z*: $[M + H]^+$ calcd for $C_{20}H_{16}NO_3$, 318.1130; found: 318.1128.



Compound 1

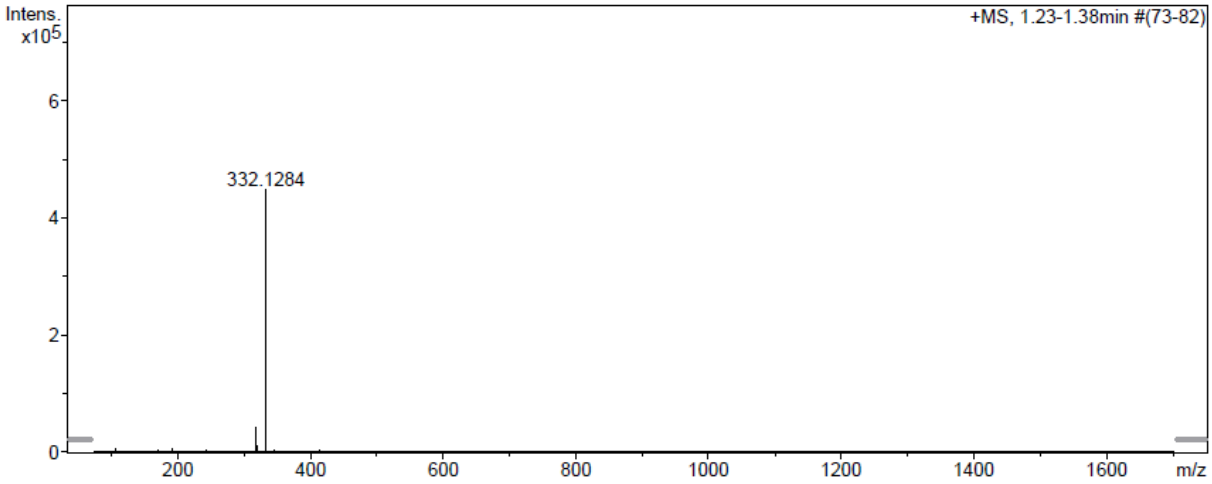


Compound **8** (79 mg, 0.25 mmol, 1 equiv.) and dimethyl sulfate (474 µL, 5 mmol, 20 equiv.) were dissolved in toluene (20 mL). The mixture was heated at 100 °C overnight. A large amount of precipitates formed. After cooling to the room temperature, the precipitate was filtered, washed with Et_2O and hexane, affording product **1** as light yellow solid (110 mg, 99% yield).

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.09 (d, J = 8.6 Hz, 1H), 8.63 (d, J = 9.2 Hz, 1H), 8.46 – 8.41 (m, 1H), 8.32 – 8.22 (m, 1H), 8.12 (d, J = 8.5 Hz, 1H), 8.00 (d, J = 9.3 Hz, 1H), 7.68 (d, J = 8.6 Hz, 1H), 7.14 – 7.09 (m, 2H), 6.31 (d, J = 9.5 Hz, 1H), 5.55 – 5.52 (m, 2H), 4.44 (s, 3H), 3.07 (s, 3H).

^{13}C NMR (125 Hz, DMSO-d_6) δ 161.2, 161.0, 160.2, 155.3, 145.4, 144.3, 138.9, 137.6, 134.2, 129.7, 128.2, 127.8, 125.6, 119.5, 113.1, 112.9, 112.9, 101.8, 68.5, 52.8, 23.0.

HRMS (ESI, pos.) *m/z*: $[M - MeSO_4]^+$ calcd for $C_{21}H_{18}NO_3$, 332.1281; found: 332.1284.



Construction of Sav-FPD mutant library for directed evolution

TOP10(DE3)_OmpA Sav-FPD construct was obtained from the previous work.[34]

Table 1. Sequence of the OmpA Sav-FPD construct for periplasmic expression (the FPD loop is highlighted in blue)

protein sequence	MKKTAIAIAVALAGFATVAQAASMTGGQQMGRDQAGITGTWYNQLGS TFIVTAGADGALTGTYESAVGNAESRYVLTGRYDSAPATDGSGTALGW TVAWKNNYRNAHSATTWSGQYVGGAEARINTQWLLTSGTTSPLSEALT KANSPAEAYKASRGAGANAWKSTLVGHDTFTKVKPSAASIDAAKKAG VNNGNPLDAVQQ
DNA sequence	ATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGC TACCGTAGCGCAGGCCGCAAGCATGACCGGTGGCCAGCAGATGGGT CGTGATCAGGCAGGTATTACCGGCACCTGGTATAATCAGCTGGGTAG CACCTTTATTGTTACCGCAGGCGCAGATGGTGCACTGACCGGTACGT ATGAAAGCGCAGTTGGTAATGCAGAAAGCCGTTATGTTCTGACAGGT CGTTATGATAGCGCACCGGCAACCGATGGTAGCGGCACCGCACTGG GTTGGACCGTTGCATGGAAAAATAACTATCGTAATGCACATAGCGCA ACCACCTGGTCAGGTCAGTATGTTGGTGGTGCAGAAGCACGCATTAA TACCCAGTGGCTGCTGACCAGCGGCACCACCAGCCCGCTGAGCGAA GCGCTGACCAAAGCGAACAGCCCGGCGGAAGCGTATAAAGCGAGCC GCGGCGCGGGCGCAAATGCCTGGAAAAGCACCCTGGTTGGTCATGA TACCTTTACCAAAGTTAAACCGAGCGCAGCAAGCATTGATGCAGCAA AAAAAGCCGGTGTGAATAATGGTAATCCGCTGGATGCAGTTCAGCA GTAATAG

Mutagenesis protocol

Four positions (S112, K121, FPD(K9) and FPD(N11)) were selected for mutagenesis using NHT codons.

For the OmpA Sav-FPD S112X construct, the above OmpA Sav-FPD sequence was selected as the template and the S112 postion was first mutated by PCR (Q5 Hot Start DNA polymerase, forward primer: 5'-CTG ACC NHT GGC ACC ACC AGC CCG CTG, reverse primer: 5'-GTG CCA DNG GTC AGC AGC CAC TGG G) (Table 2). Afterwards, the PCR mixtures were transfered to an agarose gel (1% agrose) for amplification determination. The samples with amplified bands were digested by DpnI (20U, 1 µL) for 4 h at 37 °C. Then, the PCR products (5 µL) were transformed to TOP10(DE3) competent *E. coli* cells and plated onto LB-argar plates containing kanamycin (50 µg/mL). Colonies were picked for an overnight culture. Cells were harvested and the plasmids were isolated (Macherey-Nagel NucleoSpin® Plasmid kit) and sequenced (Microsynth AG). The plasmid with the correct OmpA Sav FPD-S112 (NHT) was used as the template for the subsequent K121 mutation.

Table 2. PCR conditions

	1x MasterMix	**Final concentration**
MQ H_2O	31.75 µL	
5x Q5 reaction buffer	10 µL	
dNTPs (10 mM)	1 µL	0.2 mM
Forward Primer (10 µM)	1 µL	0.2 µM
Reverse Primer (10 µM)	1 µL	0.2 µM
Template (25 ng/µL)	1 µL	0.5 ng/ µL
DMSO	4 µL	8%
Q5 Hot Start (2 U/µL)	0.25 µL	
total	50 uL	

To prepare the OmpA Sav-FPD S112X-K121Y construct, OmpA Sav-FPD S112 (NHT) was used as the template. The PCR was performed (Q5 Hot Start polymerase, forward primer: 5'-CTG GNH TAG CAC CCT GGT TGG TCA TG, reverse primer: 5'-GGT GCT ADN CCA GGC ATT TGC GCC) using the protocol described above. After the digestion of PCR products, transformation, overnight culture, mini-prep purification and sequencing, the OmpA Sav-FPD S112 (NHT)-K121 (NHT) library was obtained and 450 clones were screened for their activity in the transfer hydrogenation of self-immolative substrate **1**.

Screening results highlighted that the OmpA Sav-FPD S112V-K121A was the best mutant and it was used as the template for the FPD loop mutagenesis. To mutate FPD(K9) and FPD(N11) simutaneously, a PCR was performed (phusion polymerase, forward primer: 5'-CCN HTG CGN HTA GCC CGG CGG AAG CGT ATA AAG, reverse primer: 5'-CTA DNC GCA DNG GTC AGC GCT TCG CTC AGC G). The final OmpA Sav-FPD(K9 (NHT)-N11 (NHT)) S112V-K121A library was obtained using the same procedure described above. Six hundred clones of this library were screened.

Directed evolution screening protocol

The plasmid solutions of the Sav-FPD mutant library were transformed to TOP10(DE3) *E. coli* competent cells and plated on LB-argar plate with kanamycin for overnight incubation at 37 °C. To express the OmpA Sav-FPD mutants in a 96-deepwell plate format, LB medium (600 µL) supplemented with kanamycin (50 µg/mL) were inoculated from a single colony from the mutant library for an overnight culture. This pre-culture (30 µL) was inoculated to a main culture (1 mL) in a modified ZYM-5052 rich medium (kanamycin 50 µg/mL) for 4 h at 37 °C. The rest of the preculture solutions were mixed with glycerol (60% in MQ H_2O) and the whole plate was stored at -80 °C as glycerol stocks. The main cultures were induced by IPTG (final concentration 50 µM) and continued shaking for another 4 h at 25 °C. Then these main cultures were treated with $Cu(gly)_2$ (20 µL from 100 mM stock, final concentration 2 mM) for 15 min.

The above main cultures were harvested by centrifugation (3200 g, 20 min) and washed with cold MOPS buffer (1 mL, 50 mM, 0.9% NaCl, pH = 7.4). After the supernatant was discarded, cell pellets were resuspended with 0.5 mL MOPS buffer which contains Ir cofactor (2 μM) on ice for 30 min for cellular uptake of cofactor. Supernatant containing excess cofactor was discarded by centrifugation (3200 g, 20 min). Then cell pellets were washed with MOPS buffer, spun down and the supernatant was discarded.

Catalysis buffer MOPS (100 μL, 0.8 M, sodium formate 2 M, pH 7.0), H_2O (80 μL) and quinolium substrate **1** (20 μL, 10 mM) were added to each well of 96-well plate and the cell pellets were resuspended. The plate was set into an incubator and shook for 16 h at 25 °C. After completion of catalysis, MeOH (400 μL) and H_2O (400 μL) were added to the reaction mixture. Then, the plate was spun down (4400 rpm, 20 min) to precipitate all cell debris. The supernatant (200 μL) was transferred to a black 96-well plate for fluorescence measurement (λ_{ex} = 322 nm, λ_{em} = 440 nm). The wells which displayed high fluorescence value were sequenced.

SDS-PAGE

To identify the expression level of periplasmic Sav, periplasmic extraction and cell lysis were performed for SDS-PAGE analysis. OD_{600} 4/mL cells were transferred to eppendorf tubes, spun down and the supernatant was discarded. The cell pellets were resuspended in 0.25 mL TSE buffer (200 mM Tris-HCl, 500 mM sucrose, 1 mM EDTA, pH 8.0). The suspensions were slowly shaken at room temperature using a Thermo mixer (23°C, 400 rpm, 10 min), and the cells were centrifuged (10,000 g, 4 °C, 2 min) and the supernatant was discarded. While sitting on ice, the cell pellets were carefully resuspended in 0.25 mL of an ice-cold $MgSO_4$ solution (5 mM). The suspension was then incubated on ice for 10 min, centrifuged (10,000 g, 4 °C, 2 min) and the supernatant was collected as the periplasmic fraction.

The remaining cell pellets (i.e. the cytoplasmic fraction) was lysed with a lysis buffer (100 uL, MOPS 0.1 M, pH 7.0, lysozyme 1 mg/mL, DNase I 0.1 mg/mL) and incubated at 37 °C for 1 h. The suspension is centrifuged and the supernatant collected as the cytoplasmic fraction. Solutions from the above two fractions were analyzed by SDS gel electrophoresis using B4F (biotin-4-fluorescein) to reveal the biotin-binding proteins. (Figure 1)

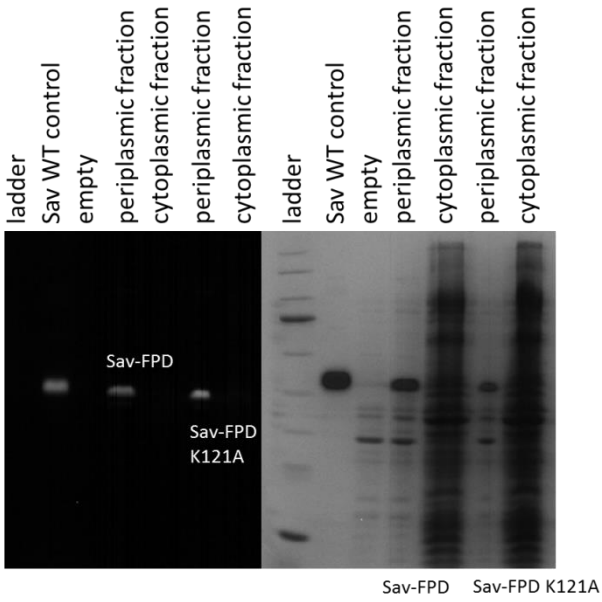


Figure 1. Example of SDS-PAGE of the Sav-FPD constructs. The acrylamide gel on the left was exposed under UV-light in the presence of B4F. The fluorescenet bands reveal the biotin-binding ability of the Sav-FPD. The comassie blue-stained gel on the right reveals the good expression of Sav-FPD in the periplasmic fraction of *E. coli* cells.

Large expression for promising Sav-FPD mutants

Following the directed evolution, four Sav-FPD mutants were selected for purification: Sav-FPD, Sav-FPD S112V-K121A, Sav-FPD(K9L-N11P) S112V-K121A, Sav-FPD(K9I-N11P) S112V-K121A.

Table 3. Sequence of the Sav-FPD construct for cytoplamic expression (the FPD loop is highlighted in blue)

protein sequence	MASMTGGQQMGRDQAGITGTWYNQLGSTFIVTAGADGALTGTYESAV GNAESRYVLTGRYDSAPATDGSGTALGWTVAWKNNYRNAHSATTWS GQYVGGAEARINTQWLLTSGTTSPLSEALTKANSPAEAYKASRGAGAN AWKSTLVGHDTFTKVKPSAASIDAAKKAGVNNGNPLDAVQQ
DNA sequence	ATGGCAAGCATGACCGGTGGCCAGCAGATGGGTCGTGATCAGGCAG GTATTACCGGCACCTGGTATAATCAGCTGGGTAGCACCTTTATTGTTA CCGCAGGCGCAGATGGTGCACTGACCGGTACGTATGAAAGCGCAGT TGGTAATGCAGAAAGCCGTTATGTTCTGACAGGTCGTTATGATAGCG CACCGGCAACCGATGGTAGCGGCACCGCACTGGGTTGGACCGTTGCA TGGAAAAATAACTATCGTAATGCACATAGCGCAACCACCTGGTCAGG TCAGTATGTTGGTGGTGCAGAAGCACGCATTAATACCCAGTGGCTGC TGACCAGCGGCACCACCAGCCCGCTGAGCGAAGCGCTGACCAAAGC GAACAGCCCGGCGGAAGCGTATAAAGCGAGCCGCGGCGCGGGCGCA AATGCCTGGAAAAGCACCCTGGTTGGTCATGATACCTTTACCAAAGT TAAACCGAGCGCAGCAAGCATTGATGCAGCAAAAAAAGCCGGTGTG AATAATGGTAATCCGCTGGATGCAGTTCAGCAGTAATAG

The OmpA tag was cleaved from the previous OmpA Sav-FPD construct to afford the Sav-FPD construct using the reported procedure.[34]

The sequences from the four mutants (selected for large scale expression and purification) were amplified by PCR (forward primer: 5'-GGG AAT TCC ATA TGG CAA GCA TGA CCG GTG GC, reverse primer: 5'-GTT AGC AGC CGG ATC TCA GTG, conditions see Table 4).

Table 4. PCR conditions for DNA amplification

	1x MasterMix	**Final concentration**
Sterilized H_2O	21 µL	
5x Q5 reaction buffer	12 µL	
5*GC enhancer	12 uL	
dNTPs (10 mM)	1.2 µL	0.2 mM
Forward Primer (10 µM)	6 µL	1 µM
Reverse Primer (10 µM)	6 µL	1 µM
Plasmid prep template (20 ng/µL)	1.5 µL	0.5 ng/µL
Q5 Hot Start (2 U/µL)	0.3 µL	
total	60 uL	

The amplified DNA was digested (NdeI and BamHI restriction enzymes) and ligated into the empty pET30 vector (New England BioLabs® T4-DNA Ligase, room temperature, 3 h, 10-fold molar ecess of inserts compared to the vector). The ligated products (3 µL) were transformed to TOP10(DE3) competent *E. coli* cells and the suspension was plated onto LB-argar plate (containing kanamycin 50 µg/mL) at 37 °C overnight. The rest of the ligated product solutions were freezen at -20 °C. Colonies were picked for overnight culture (LB medium, 37 °C). The cells were harvested and the plasmids were isolated and sequenced. The plasmids (1 µL) with the correct Sav-FPD sequence were transformed to BL21(DE3) competent cells and the suspensions were plated on LB-argar plate with kanamycin overnight at 37 °C. A single colony was picked for an overnight culture (5 mL LB medium in falcon tube) and, the next day glycerol stocks were prepared and kept at -80 °C.

Large scale expression in 3 L conical flasks using autoinduction medium

An autoinduction medium was prepared according to the reported protocol.[21] (Table 5) For each mutant, the cytoplamsic expression was performed in triplicate.

Table 5. Autoinduction medium for cytoplasmic Sav expression

	amount for 1 L solution in 3 L elenmyer flask
tryptone	10 g
yeast extract	5 g
20 · ZYP salts	50 mL
20 · ZYP sugars	50 mL
$MgSO_4$ (200 mM)	10 mL
kanamycin (50 mg/mL)	1 mL
dd H_2O	890 mL

The precultures (1 mL) was inoculated to the autoinduction medium (1 L). This main culture was incubated at 30 °C, 180 rpm for 24 h. The final OD_{600} reached to 8-10. The proteins expressed as inclusion bodies.

The cells were harvested by centrifugation (9500 rpm, 10 min, 4 °C) and the supernatant was discarded. After freezing at -20 °C overnight, the cell pellets were thawed and treated with lysis buffer (150 mL, tris-HCl 20 mM, pH 7.4, lysozyme 1 mg/mL, DNase I 0.1 mg/mL). The suspensions were shaken at room temperature for 3 h, frozen 2 h and thawed at rt. The cell lysates were dialysed dialysed in guanidium hydrochloride solution (6 M, pH 1.5) overnight at 4 °C. Then, the lysates were transfered into refolding buffer (1.5 L, MES 50 mM, pH = 6.0 NaCl 9.6 mM, KCl 0.4 mM, $MgCl_2$ 2 mM, $CaCl_2$ 2 mM, arginine 0.5 M and 0.05% polyethlene glycol 3.400) and gently stirred for 4 h at 4 °C. The resulting solutions were dialysed in an iminobiotin binding buffer (500 mM NaCl, 50 mM $NaHCO_3$, pH = 9.8) overnight at 4 °C. Finally, the cell lysates were centrifuged (10000 rpm, 2 h, 4 °C) and the supernatant was filtered and purified by AKTA prime. The eluted Sav fractions were collected, neutralised with aqueous NaOH to pH 7 and dialyzed in MQ H_2O (3 × 24 h). The clear solutions were frozen and lyophilized.

OD_{600} 4/mL cells from the main culture were transferred to Eppendorf tubes, lysed and both soluble and insoluble fractions were analyzed by SDS-PAGE. (Figure 2)

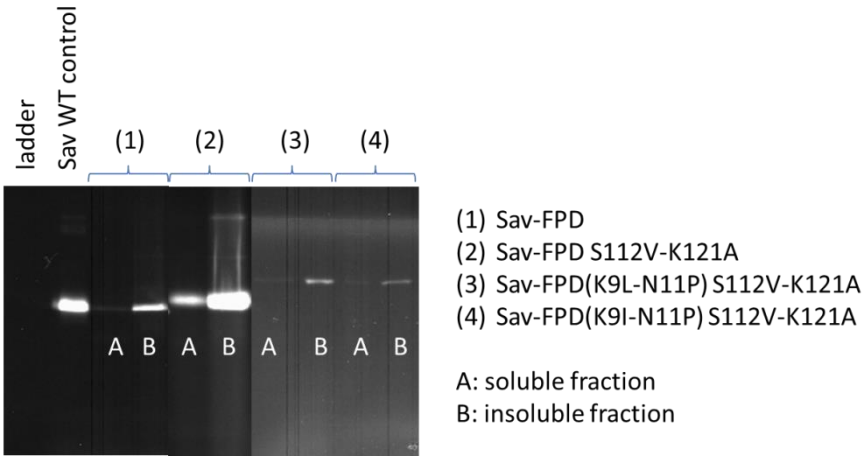


Figure 2. Overexpressed Sav-FPD mutants in the cytoplasm were visualised using B4F in the insoluble fraction under UV-light, revealing the biotin-binding activity of the inclusion bodies.

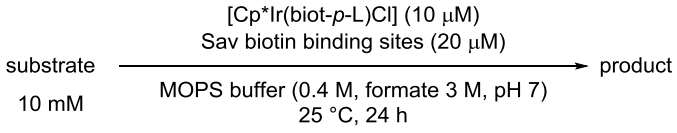
The mass of the purified Sav-FPD mutants were determined by ESI/micrOTOF mass spectroscopy. (Table 6)

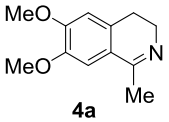
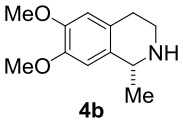
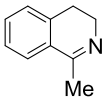
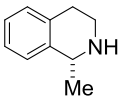
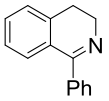
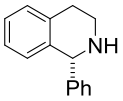
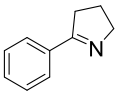
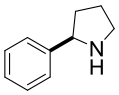
Table 6 calculated mass values and the measured mass for Sav-FPD mutants

	Calculated mass of the monomer (Da)	Measured mass (Da)
Sav-FPD	18654.40	18654.56
Sav-FPD S112V-K121A	18609.35	18609.46
Sav-FPD(K9L-N11P) S112V-K121A	18577.35	18577.45
Sav-FPD(K9I-N11P) S112V-K121A	18577.35	18577.52

General procedure of the reduction of imine 4a, 5a, 6a, 7a catalyzed by the artificial transfer hydrogenase

Table 6. Substrate scope for ATHase using mutants identified from self-immolative substrate **1**



substrate	product	mutants	TON[a]	ee (%)[a]
 **4a**	 **4b**	no protein	93 ± 0	–
		WT	84 ± 8	28 ± 3
		Sav-FPD	453 ± 24	6 ± 2
		Sav-FPD S112V-K121A	386 ± 2	-48 ± 1
		Sav-FPD (K9L-N11P) S112V-K121A	99 ± 1	-42 ± 0
		Sav-FPD (K9I-N11P) S112V-K121A	289 ± 5	-59 ± 0
 **5a**	 **5b**	no protein	315 ± 6	–
		WT	497 ± 4	49 ± 1
		Sav-FPD	676 ± 4	2 ± 0
		Sav-FPD S112V-K121A	757 ± 12	19 ± 1
		Sav-FPD (K9L-N11P) S112V-K121A	177 ± 19	25 ± 1
		Sav-FPD (K9I-N11P) S112V-K121A	600 ± 5	22 ± 1
 **6a**	 **6b**	no protein	98 ± 1	–
		WT	101 ± 7	51 ± 3
		Sav-FPD	248 ± 12	60 ± 1
		Sav-FPD S112V-K121A	355 ± 6	44 ± 1
		Sav-FPD (K9L-N11P) S112V-K121A	86 ± 3	44 ± 2
		Sav-FPD (K9I-N11P) S112V-K121A	345 ± 2	43 ± 1
 **7a**	 **7b**	no protein	61 ± 3	–
		WT	260 ± 5	78 ± 3
		Sav-FPD	342 ± 12	48 ± 1
		Sav-FPD S112V-K121A	253 ± 2	21 ± 1
		Sav-FPD (K9L-N11P) S112V-K121A	189 ± 5	33 ± 1
		Sav-FPD (K9I-N11P) S112V-K121A	436 ± 8	25 ± 0

[a] The conversions and ee were determined either by HPLC (substrates **4a**, **5a**, **6a**) or GC (for **7a**). Positive ee refers to an excess in favor of the (*R*)-amine. Negative ee refers to an excess if favour of the (*S*)-amine.

A series of stock solutions were prepared first: [Cp*Ir(biot-*p*-L)Cl] catalyst (200 μM in MQ water), streptavidin stock (400 μM biotin binding sites in MQ water), imine substrates **4a**, **5a**, **6a**, **7a** stock (400 mM in DMSO), quinolinium **1** (40 mM in DMSO) and MOPS buffer (0.8 M, formate 6 M, pH 7.0).

In a small glass reaction vial, a MOPS buffer stock solution (100 μL), the protein stock solution (10 μL), catalyst stock solution (10 μL) and MQ water (75 μL) were added and incubated at 25 °C for 10 min. Then, the imine substrate or quinolinium substrate stock solution (5 μL) was added, and the reaction mixture (200 μL in total) was shaked in a ThermoMixer at 25 °C for 24 h.

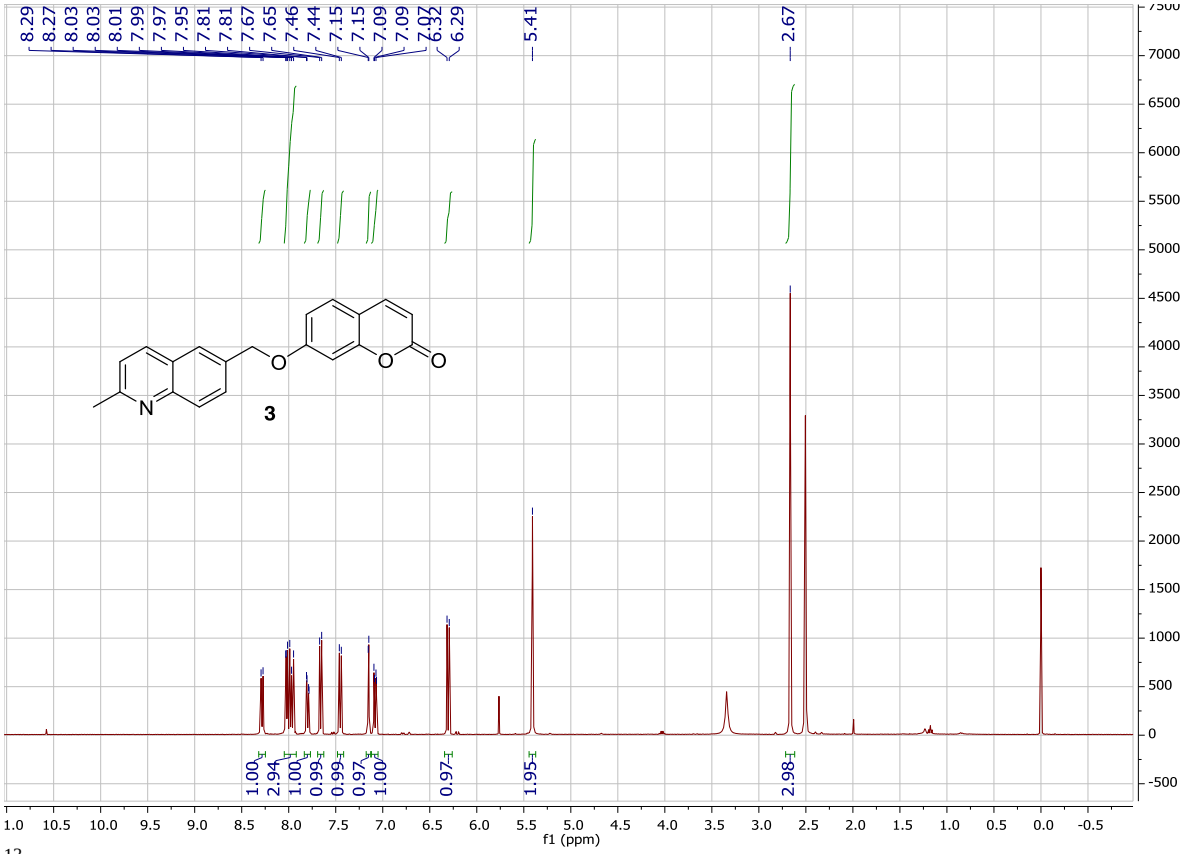
Upon completion of the reaction, the mixture was treated with aq. NaOH (100 μL, 10 M) and extracted with ethyl acetate (1 × 1 mL). The mixture was thoroughly mixed, and the organic phase was transferred to an Eppendorf tube, dried with Na_2SO_4 and centrifuged at 14'000 rpm for 10 minutes.

The supernatant (700 μL) was transferred in an HPLC vial and subjected to HPLC or GC analysis to determine the conversion and ee.

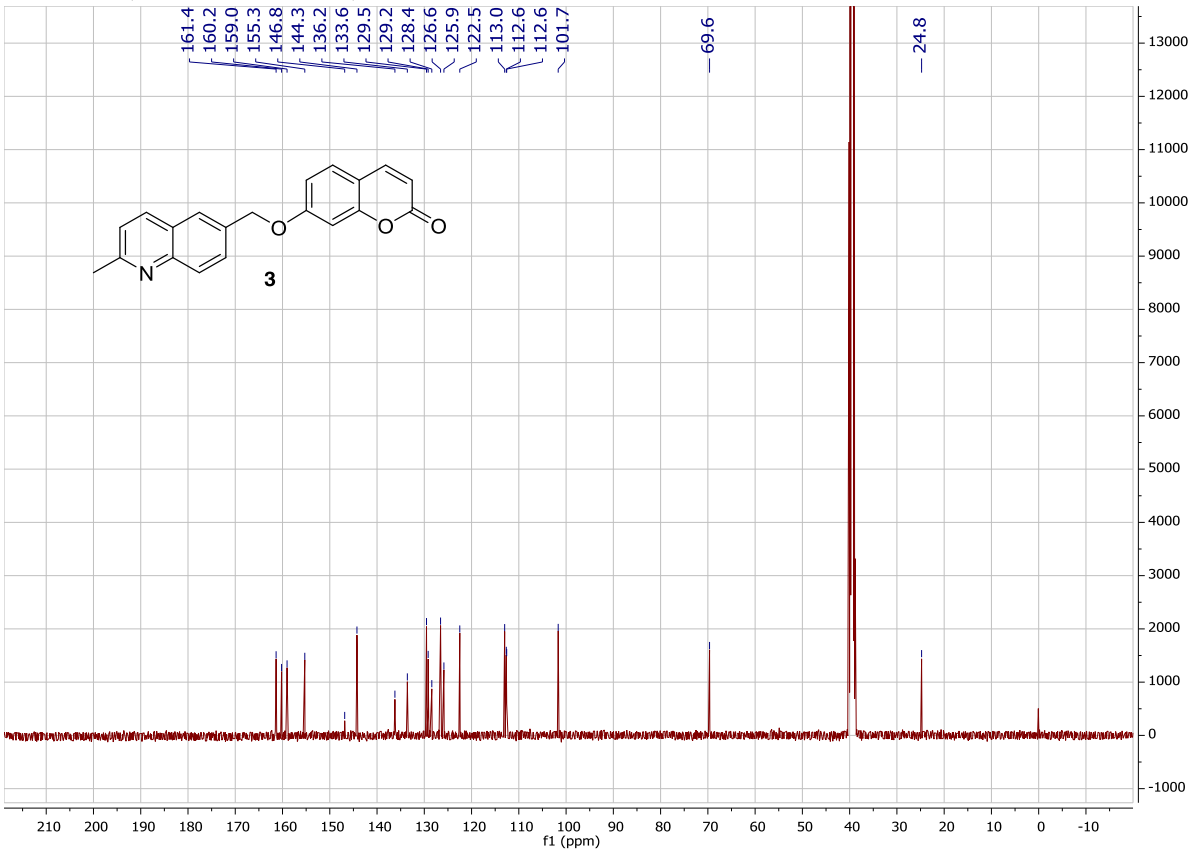
Transfer hydrogenation catalysis of imines **4a**,[34] **5a**,[45] **6a**,[20] **7a**[34] were analyzed using the reported methods.

NMR Spectra:

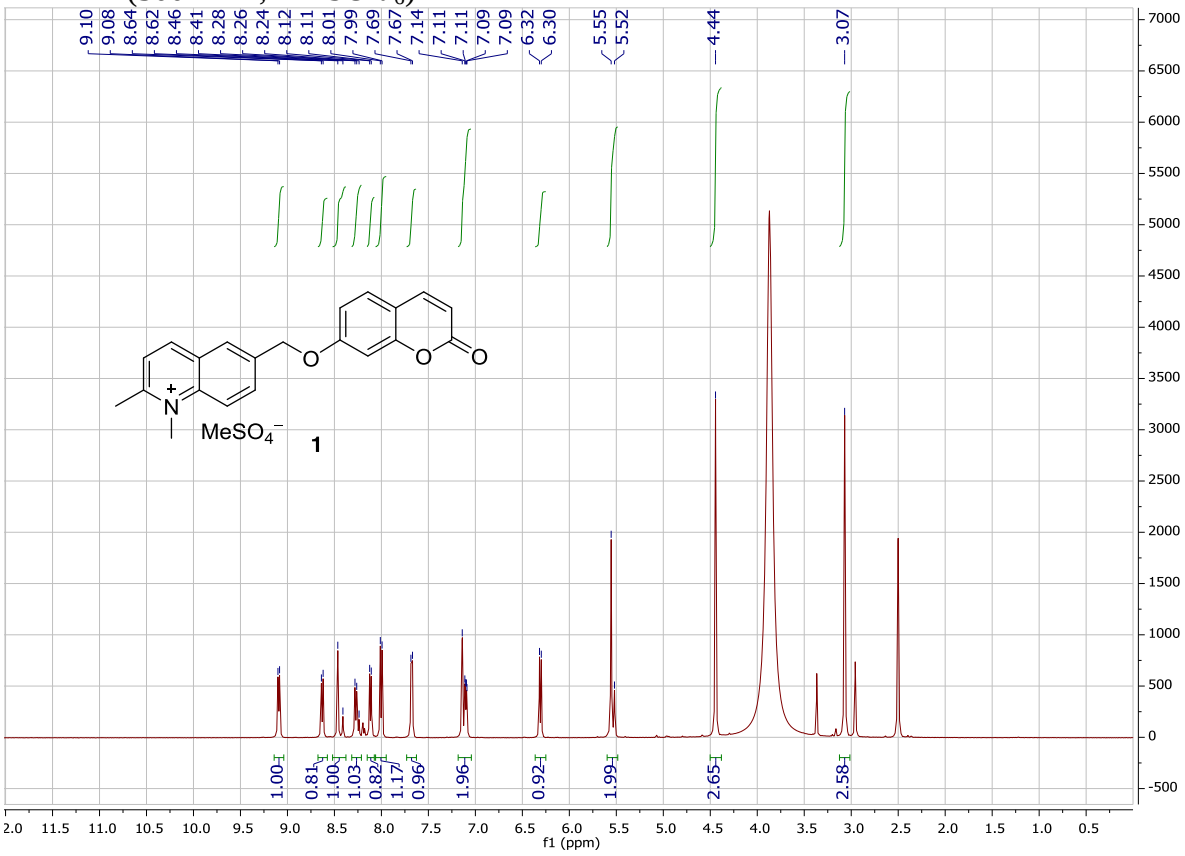
^{1}H NMR (400 MHz, DMSO-d_6)



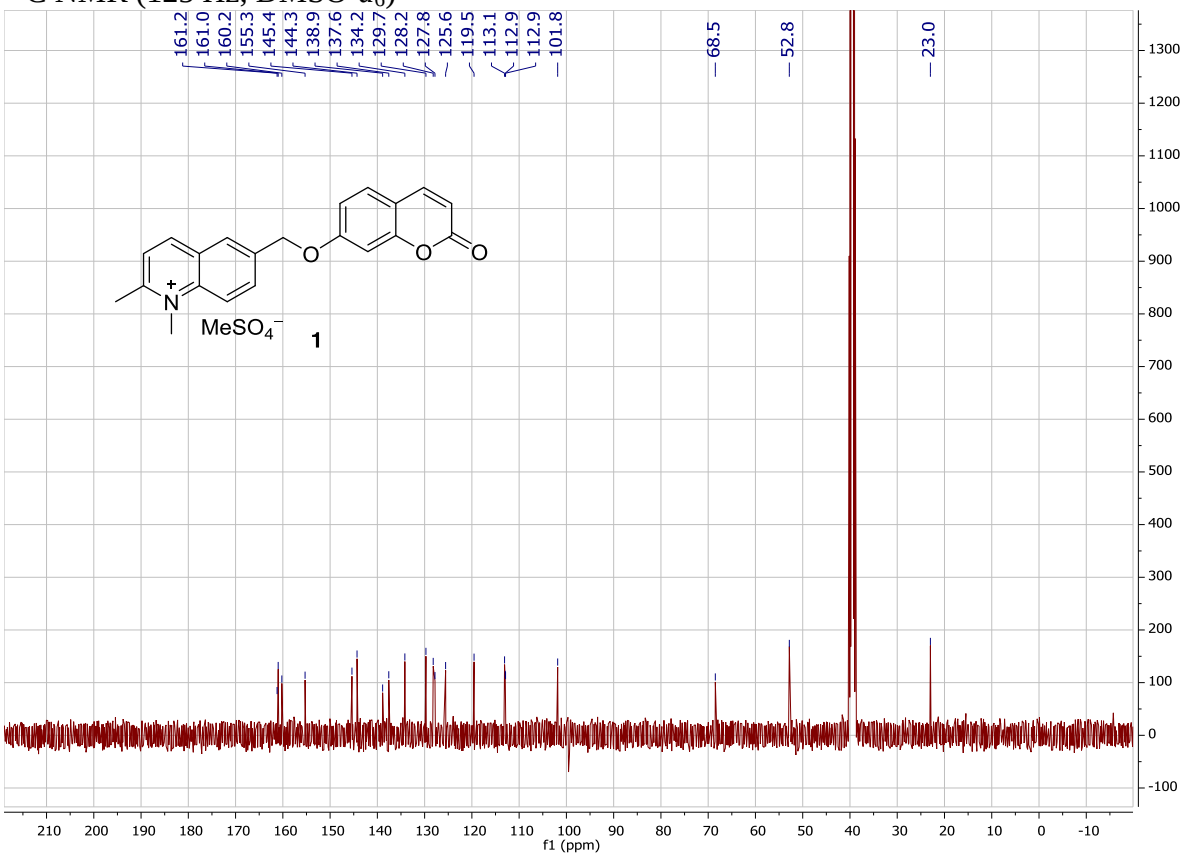
^{13}C NMR (100 Hz, DMSO-d_6)



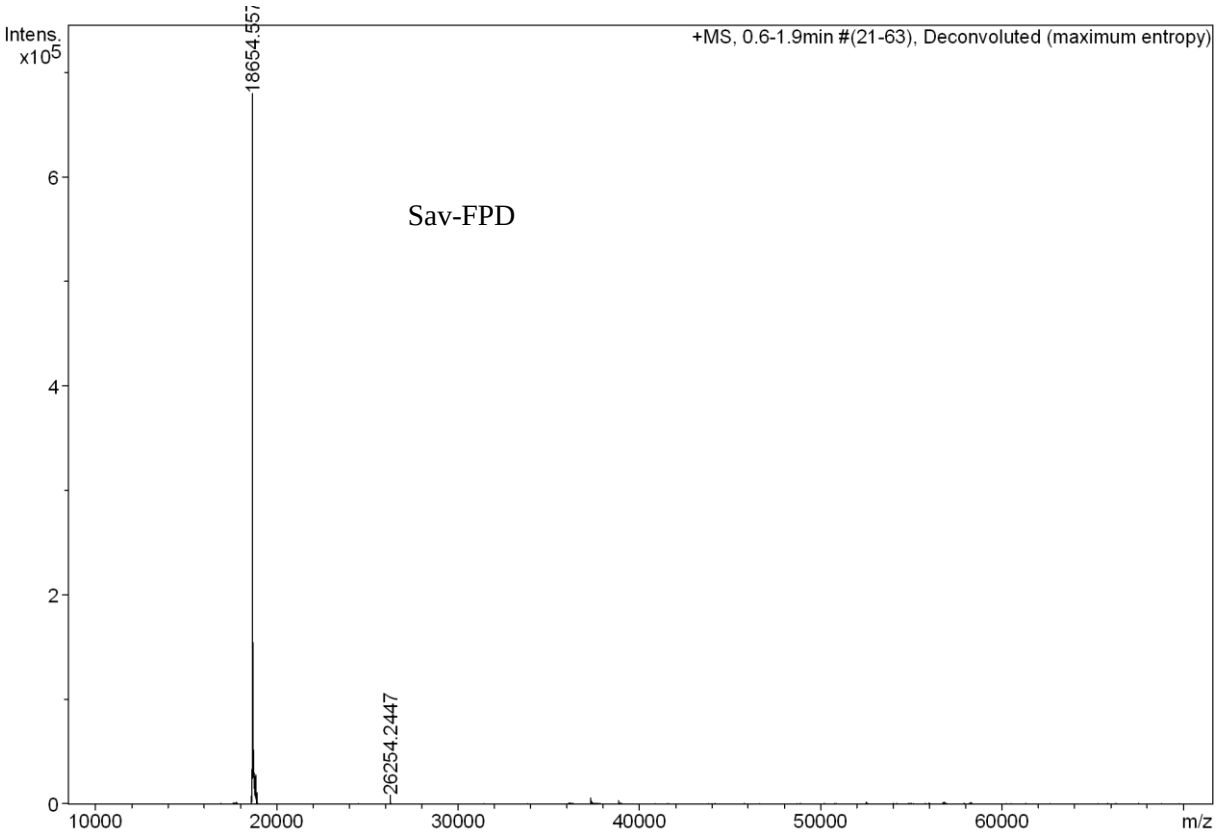
^{1}H NMR (500 MHz, DMSO-d_6)

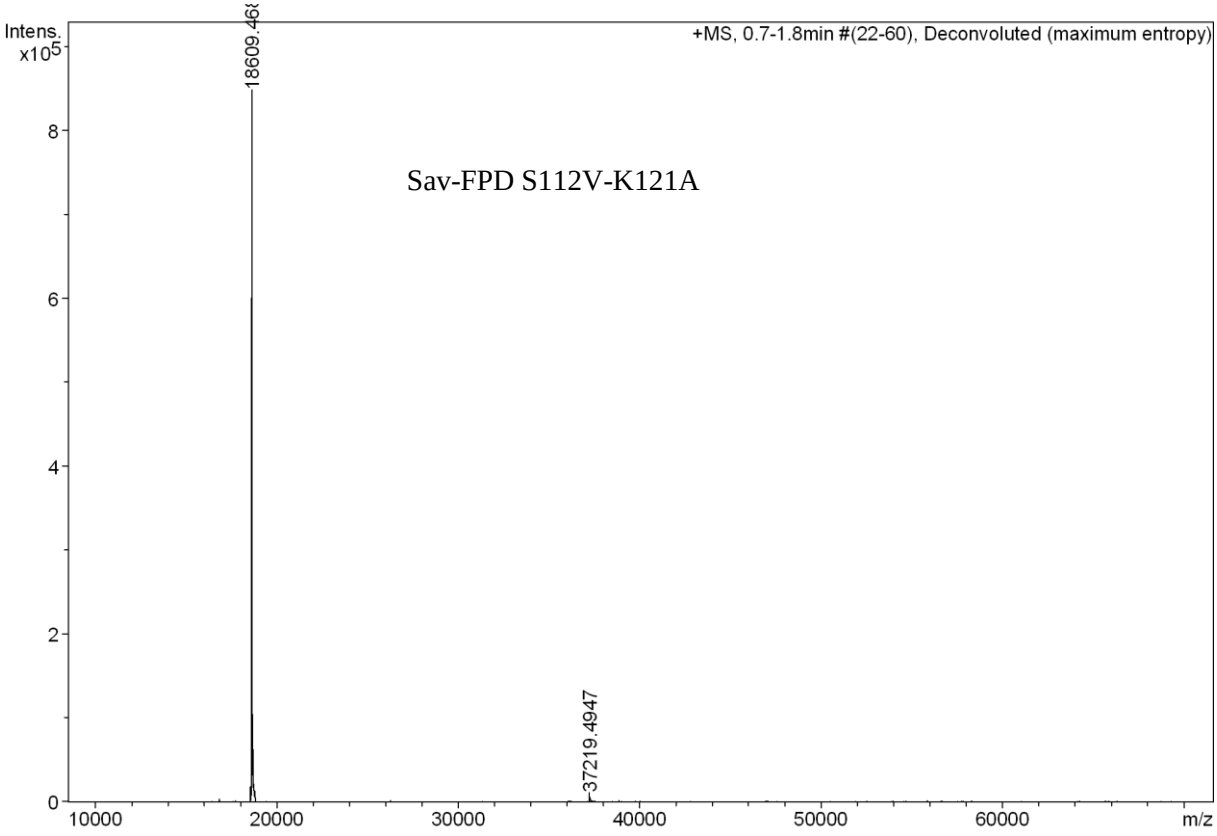


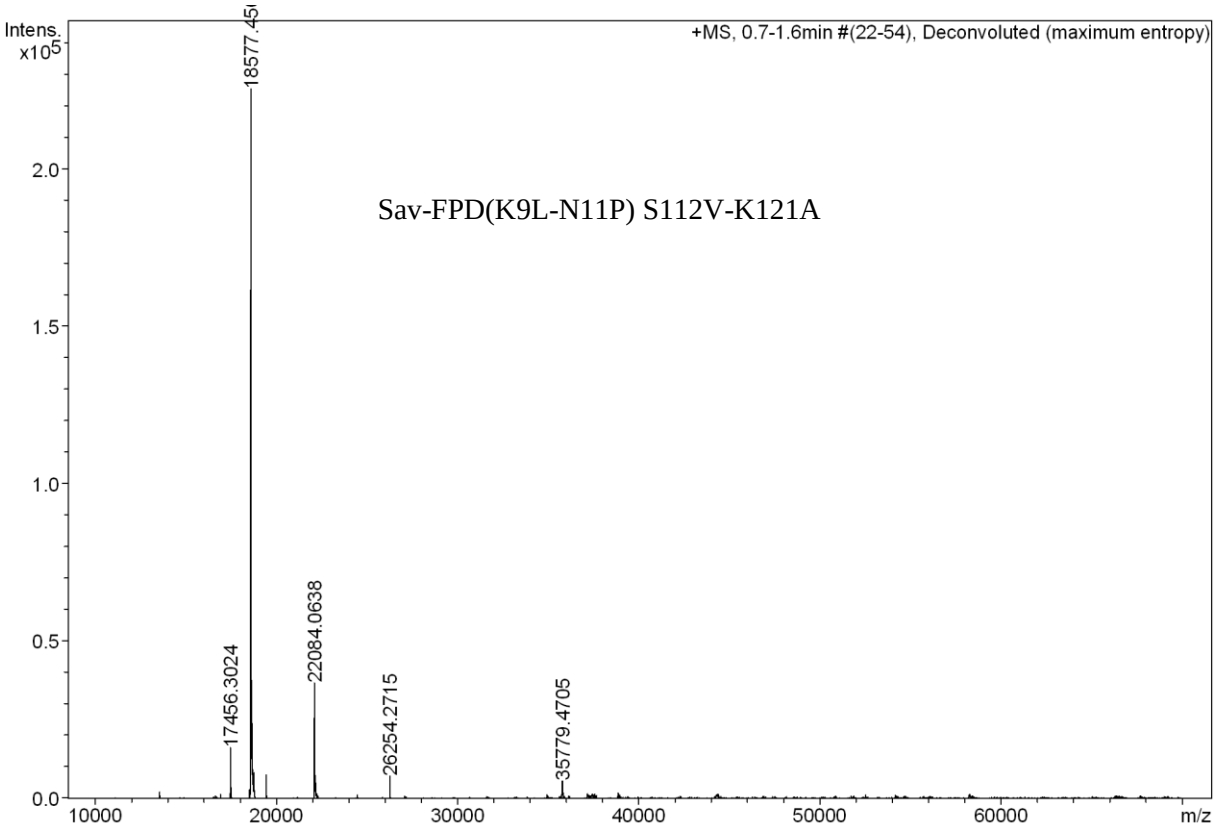
^{13}C NMR (125 Hz, DMSO-d_6)



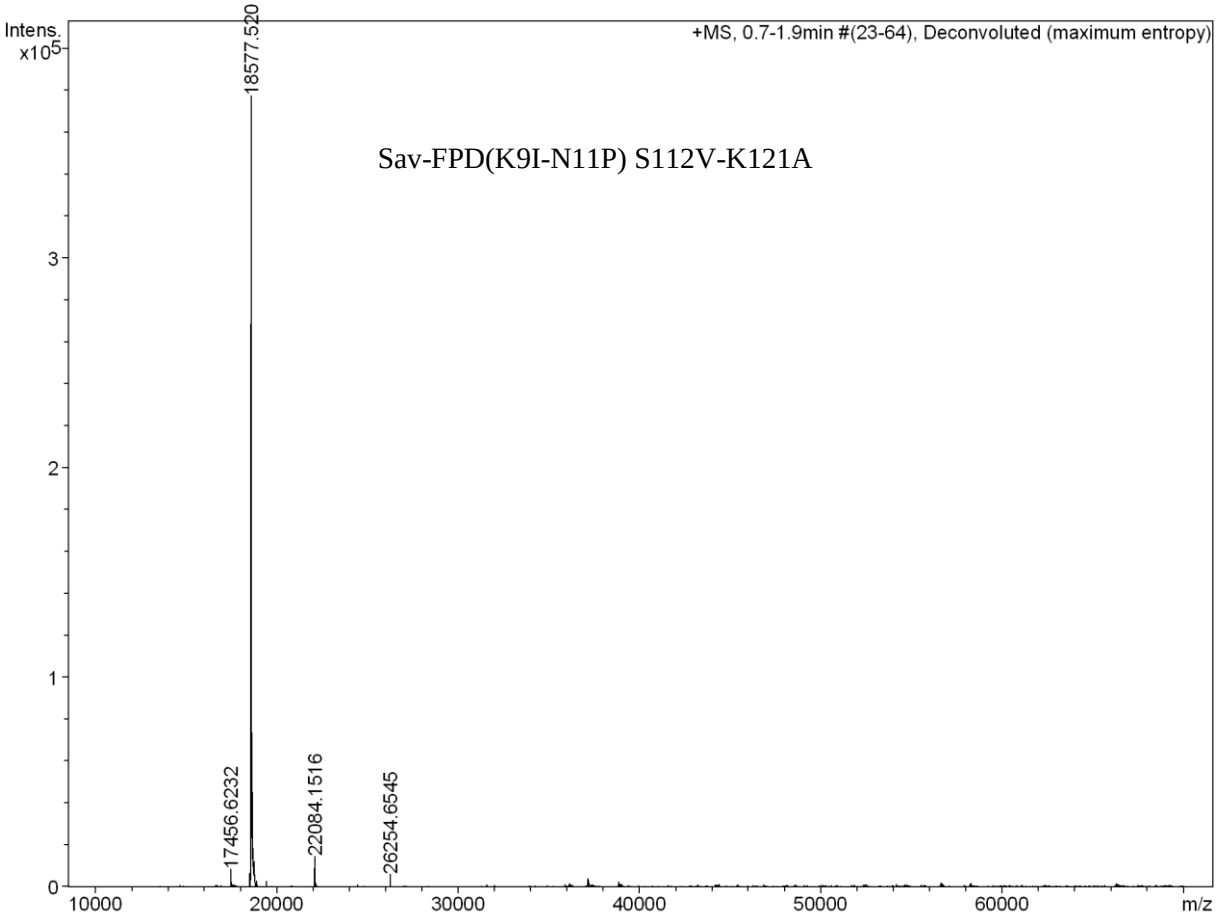
ESI/micrOTOF mass spectrum of the purified Sav-FPD mutants







Sav-FPD(K9L-N11P) S112V-K121A



Sav-FPD(K9I-N11P) S112V-K121A

4.6 References

(1) Chen, K.; Arnold, F. H. *Proc. Natl. Acad. Sci.* **1993**, 90 (12), 5618.

(2) Arnold, F. H. *Angew. Chemie., Int. Ed.* **2017**, 56, 2.

(3) Turner, N. J. *Nat. Chem. Biol.* **2009**, 5 (8), 567.

(4) Reetz, M. T.; Zonta, A.; Schimossek, K.; Liebeton, K.; Jaeger, K. E. *Angew. Chemie., Int. Ed.* **1997**, 36 (24), 2830.

(5) Wang, J.; Li, G.; Reetz, M. T. *Chem. Commun.* **2017**, 53 (28), 3916.

(6) Ilie, A.; Reetz, M. T. *Isr. J. Chem.* **2015**, 55 (1), 51.

(7) Schwizer, F.; Okamoto, Y.; Heinisch, T.; Gu, Y.; Pellizzoni, M. M.; Lebrun, V.; Reuter, R.; Köhler, V.; Lewis, J. C.; Ward, T. R. *Chem. Rev.* **2018**, 118 (1), 142.

(8) Ward, T. R. *Angew. Chemie., Int. Ed.* **2016**, 55 (48), 14909.

(9) Hyster, T. K.; Ward, T. R. *Angew. Chemie., Int. Ed.* **2016**, 55 (26), 7344.

(10) Dydio, P.; Key, H. M.; Nazarenko, A.; Rha, J. Y.-E.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. *Science* **2016**, 354 (6308), 102.

(11) Yang, H.; Swartz, A. M.; Park, H. J.; Srivastava, P.; Ellis-Guardiola, K.; Upp, D. M.; Lee, G.; Belsare, K.; Gu, Y.; Zhang, C.; Moellering, R. E.; Lewis, J. C. *Nat. Chem.* **2018**, 10 (3), 318.

(12) Srivastava, P.; Yang, H.; Ellis-Guardiola, K.; Lewis, J. C. *Nat. Commun.* **2015**, 6, 1.

(13) Obexer, R.; Godina, A.; Garrabou, X.; Mittl, P. R. E.; Baker, D.; Griffiths, A. D.; Hilvert, D. *Nat. Chem.* **2017**, 9 (1), 50.

(14) Song, W. J.; Yu, J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2017**, 139 (46), 16772.

(15) Key, H. M.; Dydio, P.; Clark, D. S.; Hartwig, J. F. *Nature* **2016**, 534 (7608), 534.

(16) Brandenberg, O. F.; Fasan, R.; Arnold, F. H. *Curr. Opin. Biotechnol.* **2017**, 47, 102.

(17) Wilson, M. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**, 100 (1), 306.

(18) Grimm, A. R.; Sauer, D. F.; Polen, T.; Zhu, L.; Hayashi, T.; Okuda, J.; Schwaneberg, U. *ACS Catal.* **2018**, 8 (3), 2611.

(19) Wilson, Y. M.; Dürrenberger, M.; Nogueira, E. S.; Ward, T. R. *J. Am. Chem. Soc.* **2014**, 136 (25), 8928.

(20) Hestericová, M.; Heinisch, T.; Alonso-Cotchico, L.; Maréchal, J. D.; Vidossich, P.; Ward, T. R. *Angew. Chemie., Int. Ed.* **2018**, 57 (7), 1863.

(21) Mallin, H.; Hestericová, M.; Reuter R.; Ward, T. R. *Nat. Protoc.* **2016**, 11 (5), 835.

(22) Alouane, A.; Labruère, R.; Le Saux, T.; Schmidt, F.; Jullien, L. *Angew. Chemie., Int. Ed.* **2015**, 54 (26), 7492.

(23) Wang, Q.; Franz, K. J. *Acc. Chem. Res.* **2016**, 49 (11), 2468.

(24) Shamis, M.; Lode, H. N.; Shabat, D. *J. Am. Chem. Soc.* **2004**, 126 (6), 1726.

(25) Shabat, D.; Amir, R. J.; Gopin, A.; Pessah, N.; Shamis, M. *Chem. - Eur. J.* **2004**, 10 (11), 2626.

(26) Srinivasarao, M.; Low, P. S. *Chem. Rev.* **2017**, 117 (19), 12133.

(27) Hettiarachchi, S. U.; Prasai, B.; McCarley, R. L. *J. Am. Chem. Soc.* **2014**, 136 (21), 7575.

(28) Gnaim, S.; Shabat, D. *Acc. Chem. Res.* **2014**, 47 (10), 2970.

(29) Genz, M.; Köhler, V.; Krauss, M.; Singer, D.; Hoffmann, R.; Ward, T. R.; Sträter, N. *ChemCatChem* **2014**, 6 (3), 736.

(30) Reetz, M. T.; Peyralans, J. J.-P.; Maichele, A.; Fu, Y.; Maywald, M. *Chem. Commun.* **2006**, 41, 4318.

(31) Chevalley, A.; Salmain, M. *Chem. Commun.* **2012**, 48 (98), 11984.

(32) Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. *Angew. Chemie., Int. Ed.* **2011**, 50 (13), 3026.

(33) Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; LeTrong, I.; Stenkamp, R. E.; Ward, T. R. *Angew. Chemie., Int. Ed.* **2008**, 47 (8), 1400.

(34) Pellizzoni, M. M.; Schwizer, F.; Wood, C. W.; Sabatino, V.; Cotelle, Y.; Matile, S.; Woolfson, D. N.; Ward, T. R. *ACS Catal.* **2018**, 8 (2), 1476.

(35) Pordea, A.; Creusa, M.; Panek, J.; Duboc, C.; Mathis, D.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2008**, 130 (25), 8085.

(36) Ward, T. R. *Acc. Chem. Res.* **2011**, 44 (1), 47.

(37) Heinisch, T.; Ward, T. R. *Acc. Chem. Res.* **2016**, 49 (9), 1711.

(38) Eiben, C. B.; Siegel, J. B.; Bale, J. B.; Cooper, S.; Khatib, F.; Shen, B. W.; Players, F.; Stoddard, B. L.; Popovic, Z.; Baker, D. *Nat. Biotechnol.* **2012**, 30 (2), 190.

(39) Robles, V. M.; Dürrenberger, M.; Heinisch, T.; Lledós, A.; Schirmer, T.; Ward, T. R.; Maréchal, J. D. *J. Am. Chem. Soc.* **2014**, 136 (44), 15676.

(40) Jeschek, M.; Reuter, R.; Heinisch, T.; Trindler, C.; Klehr, J.; Panke, S.; Ward, T. R. *Nature* **2016**, 537 (7622), 661.

(41) Jeschek, M.; Panke, S.; Ward, T. R. *Methods Enzymol.*, **2016**, 580, 539.

(42) Kachur, A. V; Koch, C. J.; Biaglow, *J. E. Free Radic. Res.* **1998**, 28, 259.

(43) Ngamchuea, K.; Batchelor-McAuley, C.; Compton, R. G. *Chem. - Eur. J.* **2016**, 22 (44), 15937.

(44) Schwizer, F.; Köhler, V.; Dürrenberger, M.; Knörr, L.; Ward, T. R. *ACS Catal.* **2013**, 3 (8), 1752.

(45) Köhler, V.; Wilson, Y.; Dürrenberger, M.; Ghislieri, D.; Churakova, E.; Quinto, T.; Knörr, L.; Häussinger, D.; Hollmann, F.; Turner, N. J.; Ward, T. R. *Nat. chem.* **2013**, 5, 93.

Chapter 5 | Conclusion and outlook

Three ArMs, combining complexes of either Ru, Rh or Ir with the protein scaffolds human carbonic anhydrase II or streptavidin were explored. Good activities of these ArMs were observed for non-natural reactions: namely ring closing metathesis, carbene transfer reactions and transfer hydrogenation.

The study presented in chapter 2 demonstrates robust activity of a ruthenium-based artificial metathesase using the human carbonic anhydrase II scaffold with low catalyst concentration and physiological conditions. The hCA II mutants displayed binding affinities for the Ru-complex in the nanomolar range. This ArM would be suitable for further applications in cascade reactions e.g. for the synthesis of cyclic alkenes, in which natural enzymes involved in the cascade require a pH-neutral environment. So far no protein-acceleration has been observed for this system. This might be overcome by further substrate and scaffold engineering efforts. Directed evolution of hCA II for the task at hand would be an attractive way to search the sequence landscape for improved activity. In particular the combination with computational design for informed choices on randomization sites should be considered.

Chapter 3 reports on the successful synthesis of various biotinylated dirhodium complexes and the good activity of these complexes embedded in streptavidin mutants. The dirhodium ArMs catalysed intermolecular cyclopropanation and C-H insertion with high efficiency under neutral conditions. The ArMs assembled in the periplasm of *E. coli* cells showed good activity for cyclopropanation. A modelled structure of a dirhodium complex within Sav suggested that the dirhodium metal centers are too distant from any amino acid residues from the protein scaffold, thereby providing an rationale why no enantioselectivity for the tested cyclopropanation and C-H insertion was observed. To simulate the traditional chiral dirhodium complexes carrying bulky carboxylate ligands, additional motifs could be added close to the dirhodium center, such as loops inserted at different locations around the putative active site.

Chapter 4 summarizes the implementation of a streamlined screening protocol for the evaluation of loop mutants of an artificial transfer hydrogenase. The identified, purified loop mutants outperformed Sav WT in combination with the iridium-cofactor in the transfer hydrogenation of imines up to fivefold in terms of turnover numbers, although the enantioselectivity was not improved significantly. A self-immolation strategy was successfully used to analyse transfer hydrogenation catalysis results by fluorescence, greatly increasing the screening thoughput.

Directed evolution is a powerful method to engineer enzymes for a particular application. Beneficial mutations are accumulated over multiple generations and enhance the enzymes' performance. As the situation is even more complicated for the engineering of ArMs by directed evolution, some issues should be considered before a project of directed evolution of artificial metalloenzymes is undertaken.

(1) It would be advantageous if protein concentrations could be kept equal among members of a mutant library to not only match catalyst concentration, but also cofactor/host-protein ratio. (2) A suitable reaction and reaction conditions have to be identified which are sensitive to subtle changes of the catalyst structure. In a typical protocol mutations of residues near the active site would be targeted first, their choice often inspired by modelling studies. (3) It should be considered whether the system is indeed suitable for directed evolution protocols. E.g. if the catalytic center is projected outside of the protein scaffold, as hypothesized for the dirhodium constructs presented in chapter 3, high-throughput screening efforts without substantial protein engineering (e.g. the introduction of additional loops) is unlikely to yield improved mutants.